

# NORMAL FORMS FOR NON-UNIFORM CONTRACTIONS

BORIS KALININ AND VICTORIA SADOVSKAYA\*

**ABSTRACT.** Let  $f$  be a measure-preserving transformation of a Lebesgue space  $(X, \mu)$  and let  $\mathcal{F}$  be its extension to a bundle  $\mathcal{E} = X \times \mathbb{R}^m$  by smooth fiber maps  $\mathcal{F}_x : \mathcal{E}_x \rightarrow \mathcal{E}_{f_x}$  so that the derivative of  $\mathcal{F}$  at the zero section has negative Lyapunov exponents. We construct a measurable system of smooth coordinate changes  $\mathcal{H}_x$  on  $\mathcal{E}_x$  for  $\mu$ -a.e.  $x$  so that the maps  $\mathcal{P}_x = \mathcal{H}_{f_x} \circ \mathcal{F}_x \circ \mathcal{H}_x^{-1}$  are sub-resonance polynomials in a finite dimensional Lie group. Our construction shows that such  $\mathcal{H}_x$  and  $\mathcal{P}_x$  are unique up to a sub-resonance polynomial. As a consequence, we obtain the centralizer theorem that the coordinate change  $\mathcal{H}$  also conjugates any commuting extension to a polynomial extension of the same type. We apply our results to a measure-preserving diffeomorphism  $f$  with a non-uniformly contracting invariant foliation  $W$ . We construct a measurable system of smooth coordinate changes  $\mathcal{H}_x : W_x \rightarrow T_x W$  such that the maps  $\mathcal{H}_{f_x} \circ f \circ \mathcal{H}_x^{-1}$  are polynomials of sub-resonance type. Moreover, we show that for almost every leaf the coordinate changes exist at each point on the leaf and give a coherent atlas with transition maps in a finite dimensional Lie group.

## 1. INTRODUCTION

The theory of normal forms for smooth maps originated in the works of Poincare and Sternberg [St57] and normal forms at fixed points and invariant manifolds have been extensively studied [BKo]. More recently, non-stationary normal form theory was developed in the context of a diffeomorphism  $f$  contracting a foliation  $W$ . The goal is to obtain a family of diffeomorphisms  $\mathcal{H}_x : W_x \rightarrow T_x W$  such that the maps

$$(1.1) \quad \tilde{f}_x = \mathcal{H}_{f_x} \circ f \circ \mathcal{H}_x^{-1} : T_x W \rightarrow T_{f_x} W$$

are as simple as possible, for example linear maps or polynomial maps in a finite dimensional Lie group. Such a map  $\tilde{f}_x$  is called a normal form of  $f$  on  $W_x$ .

The non-stationary normal form theory started with the linearization along one-dimensional foliations obtained by Katok and Lewis [KtL91]. In a more general setting of contractions with narrow band spectrum, it was developed by Guysinsky and Katok [GKt98, G02], and a differential geometric point of view was presented by Feres [Fe04]. For the linearization, further results were obtained by the second author in [S05] and it was shown in [KS06] that the coordinates  $\mathcal{H}_x$  give a consistent affine atlas on each leaf of  $W$ . In [KS15] we extended these results to the general narrow band case. More precisely, we gave a construction of  $\mathcal{H}_x$  that depend smoothly on  $x$  along the leaves

---

*Date:* May 13, 2016.

\* Supported in part by NSF grant DMS-1301693.

and proved that they define an atlas with transition maps in a finite dimensional Lie group. Non-stationary normal forms were used extensively in the study of rigidity of uniformly hyperbolic dynamical systems and group actions, see for example [KtSp97, KS03, KS06, F07, FFH10, GoKS11, FKSp11].

To obtain applications for non-uniformly hyperbolic systems and actions, one needs a similar theory of non-stationary normal forms for non-uniform contractions. The existence and centralizer theorems were stated without proof in [KKt01] along with a program of potential applications. The theory, however, was not developed for quite a while. The linearization of a  $C^{1+\alpha}$  diffeomorphism along a one-dimensional non-uniformly contracting foliation was constructed in [KKt07] and used in the study of measure rigidity in [KKt07, KKtR11]. Similar results for higher dimensional foliations with pinched exponents were obtained by Katok and Rodriguez Hertz in [KtR15]. The existence of  $\mathcal{H}_x$  for a general  $C^\infty$  extension was proved by Li and Lu [LL05].

In this paper we develop the theory of non-stationary polynomial normal forms for smooth extensions of measure preserving transformations by non-uniform contractions, described in the beginning of Section 2. This is a convenient general setting for the construction. The foliation setting reduces to it by locally identifying the leaf  $W_x$  with its tangent space  $\mathcal{E}_x = T_x W$  and viewing  $\mathcal{F}_x = f|_{W_x} : \mathcal{E}_x \rightarrow \mathcal{E}_{fx}$  as an extension of the base system  $f : \mathcal{M} \rightarrow \mathcal{M}$  by smooth maps on the bundle  $\mathcal{E} = TW$ . The base system can then be viewed as just a measure preserving one. In the extension setting, the map  $\mathcal{H}_x$  is a coordinate change on  $\mathcal{E}_x$  and we denote

$$\mathcal{P}_x = \mathcal{H}_{fx} \circ \mathcal{F}_x \circ \mathcal{H}_x^{-1} : \mathcal{E}_x \rightarrow \mathcal{E}_{fx}.$$

In Theorem 2.3 we construct coordinate changes  $\mathcal{H}_x$  for  $\mu$  almost every  $x$  so that  $\mathcal{P}_x$  is a sub-resonance polynomial. For any regularity of  $\mathcal{F}$  above the critical level, we obtain  $\mathcal{H}$  in the same regularity class.

Our construction allows us to describe the exact extent of non-uniqueness in  $\mathcal{H}_x$  and  $\mathcal{P}_x$ . Essentially, they are defined up to a sub-resonance polynomial. As a consequence of this, we obtain the centralizer theorem that the coordinate change  $\mathcal{H}$  also conjugates any commuting extension to a normal form of the same type. We just learned of similar results in differential geometric formulations by Melnick [M16]. The approach in [M16] is different from ours and it relies on ergodic theorems for higher jets of  $\mathcal{F}_x$ . Our results assume only temperedness of the higher derivatives of  $\mathcal{F}_x$  rather than certain integrability required in [M16]. This allows us to obtain applications to the foliation setting without any assumptions on transverse regularity of the foliation.

In particular, we consider a diffeomorphism  $f$  which preserves an ergodic measure with some negative Lyapunov exponents and take  $W$  to be any strong part of the stable foliation. In this setting Theorem 2.5 gives sub-resonance normal forms for  $f$  along the leaves of  $W$ . Moreover, we show that for almost every leaf the normal form coordinates  $\mathcal{H}_x$  exist at each point on the leaf and give a coherent atlas with transition maps in a finite dimensional Lie group  $G$  determined by sub-resonance polynomials. This yields an invariant structure of a  $G$  homogeneous space on almost every leaf.

We expect these results to be useful in the study of non-uniformly hyperbolic systems and group actions.

## 2. STATEMENTS OF RESULTS

**Assumptions 2.1.** *In this paper,*

*$(X, \mu)$  is a Lebesgue probability space,*

*$f : X \rightarrow X$  is an invertible ergodic measure-preserving transformation of  $(X, \mu)$ ,*

*$\mathcal{E} = X \times \mathbb{R}^m$  is a finite dimensional vector bundle over  $X$ ,*

*$\mathcal{V}$  is a neighborhood of the zero section in  $\mathcal{E}$ ,*

*$\mathcal{F} : \mathcal{V} \rightarrow \mathcal{E}$  is a measurable extension of  $f$  that preserves the zero section,*

*$F : \mathcal{E} \rightarrow \mathcal{E}$  is the derivative of  $\mathcal{F}$  at zero section,  $F_x = D_0 \mathcal{F}_x : \mathcal{E}_x \rightarrow \mathcal{E}_{fx}$ ,*

*$F$  and  $F^{-1}$  exist and satisfy  $\log \|F_x\| \in L^1(X, \mu)$  and  $\log \|F_x^{-1}\| \in L^1(X, \mu)$ ,*

*and the Lyapunov exponents of  $F$  are negative:  $\chi_1 < \dots < \chi_\ell < 0$ .*

**Sub-resonance polynomials.** Let  $\chi_1 < \dots < \chi_\ell < 0$  be the distinct Lyapunov exponents of  $F$  and let  $\mathcal{E}_x = \mathcal{E}_x^1 \oplus \dots \oplus \mathcal{E}_x^\ell$  be the splitting of  $\mathcal{E}_x$  for  $x \in \Lambda$  into the Lyapunov subspaces given by the Multiplicative Ergodic Theorem 3.1.

We say that a map between vector spaces is *polynomial* if each component is given by a polynomial in some, and hence every, bases. We consider a polynomial map  $P : \mathcal{E}_x \rightarrow \mathcal{E}_y$  with  $P(0_x) = 0_y$  and split it into components  $(P_1(t), \dots, P_\ell(t))$ , where  $P_i : \mathcal{E}_x \rightarrow \mathcal{E}_y^i$ . Each  $P_i$  can be written uniquely as a linear combination of polynomials of specific homogeneous types: we say that  $Q : \mathcal{E}_x \rightarrow \mathcal{E}_y^i$  has homogeneous type  $s = (s_1, \dots, s_\ell)$  if for any real numbers  $a_1, \dots, a_\ell$  and vectors  $t_j \in \mathcal{E}_x^j$ ,  $j = 1, \dots, \ell$ , we have

$$(2.1) \quad Q(a_1 t_1 + \dots + a_\ell t_\ell) = a_1^{s_1} \dots a_\ell^{s_\ell} Q(t_1 + \dots + t_\ell).$$

**Definition 2.2.** *We say that a polynomial map  $P : \mathcal{E}_x \rightarrow \mathcal{E}_y$  is sub-resonance if each component  $P_i$  has only terms of homogeneous types  $s = (s_1, \dots, s_\ell)$  satisfying sub-resonance relations*

$$(2.2) \quad \chi_i \leq \sum s_j \chi_j, \quad \text{where } s_1, \dots, s_\ell \text{ are non-negative integers.}$$

*We denote by  $\mathcal{S}_{x,y}$  the space of all sub-resonance polynomial maps from  $\mathcal{E}_x$  to  $\mathcal{E}_y$ .*

Clearly, for any sub-resonance relation we have  $s_j = 0$  for  $j < i$  and  $\sum s_j \leq \chi_1 / \chi_\ell$ . It follows that sub-resonance polynomial maps have degree at most

$$(2.3) \quad d = d(\chi) = \lfloor \chi_1 / \chi_\ell \rfloor.$$

Sub-resonance polynomial maps  $P : \mathcal{E}_x \rightarrow \mathcal{E}_x$  with  $P(0) = 0$  with invertible derivative at the origin form a group with respect to composition [GKt98]. We will denote this finite-dimensional Lie group by  $G_x^\chi$ . All groups  $G_x^\chi$  are isomorphic, moreover, any map  $P \in \mathcal{S}_{x,y}$  with  $P(0_x) = 0_y$  and invertible derivative at  $0_x$  induces an isomorphism between  $G_x^\chi$  and  $G_y^\chi$  by conjugation.

We denote by  $B_{x,\sigma(x)}$  the closed ball of radius  $\sigma(x)$  centered at  $0 \in \mathcal{E}_x$ . For  $N \geq 1$  and  $0 < \alpha \leq 1$  we denote by  $C^{N,\alpha}(B_{x,\sigma(x)}) = C^{N,\alpha}(B_{x,\sigma(x)}, \mathcal{E}_x)$  the space of functions

from  $B_{x,\sigma(x)}$  to  $\mathcal{E}_x$  with continuous derivatives up to order  $N \geq 1$  on  $B_{x,\sigma(x)}$  and with  $N^{th}$  derivative satisfying  $\alpha$ -Hölder condition at 0:

$$(2.4) \quad \|D^{(N)}R\|_\alpha = \sup \{ \|D_t^{(N)}R - D_0^{(N)}R\| \cdot \|t\|^{-\alpha} : 0 \neq t \in B_{x,\sigma(x)} \} < \infty.$$

We call  $\|D^{(N)}R\|_\alpha$  the  $\alpha$ -Hölder constant of  $D^{(N)}R$  at 0. We equip the space  $C^{N,\alpha}(B_{x,\sigma(x)})$  with the norm

$$(2.5) \quad \|R\|_{C^{N,\alpha}(B_{x,\sigma(x)})} = \max \{ \|R\|_0, \|D^{(1)}R\|_0, \dots, \|D^{(N)}R\|_0, \|D^{(N)}R\|_\alpha \},$$

where  $\|D^{(k)}R\|_0 = \sup \{ \|D_t^{(k)}R\| : t \in B_{x,\sigma(x)} \}$ .

We say that a non-negative real-valued function  $K$  on  $X$  is  $\varepsilon$ -tempered at  $x$  if

$$(2.6) \quad \sup \{ K(f^n x) e^{-\varepsilon n} : n \in \mathbb{N} \} < \infty,$$

and that  $K$  is  $\varepsilon$ -tempered on a set if it is  $\varepsilon$ -tempered at each of its points.

We consider an extension  $\mathcal{F}$  satisfying the Assumptions 2.1 and denote by  $\Lambda$  the set of regular points and by  $\chi_1 < \dots < \chi_\ell < 0$  the Lyapunov exponents of  $F$  given by the Multiplicative Ergodic Theorem 3.1. For  $N$  and  $\alpha$  as above we define

$$(2.7) \quad \kappa = 1 + 3/\alpha \text{ if } N = 1 \text{ and } \kappa = 4 \text{ if } N \geq 2.$$

If  $N \geq 2$  we allow  $\alpha = 0$ , in which case we understand  $C^{N,\alpha}$  as  $C^N$ .

**Theorem 2.3** (Normal forms for non-uniformly contracting extensions).

Let  $\mathcal{F}$  be an extension of  $f$  satisfying Assumptions 2.1. Suppose that

$$(2.8) \quad N \geq 1, \quad 0 \leq \alpha \leq 1 \quad \text{and} \quad N + \alpha > \chi_1/\chi_\ell.$$

Then there exist positive constants  $L = L(N, \alpha)$  and  $\varepsilon_* = \varepsilon_*(N, \alpha, \chi_1, \dots, \chi_\ell)$  so that for any  $0 < \varepsilon \leq \varepsilon_*$  the following holds.

If there exists a positive measurable function  $\sigma : \Lambda \rightarrow \mathbb{R}$  so that  $1/\sigma$  is  $\varepsilon$ -tempered on  $\Lambda$  and  $\mathcal{F}_x$  is  $C^{N,\alpha}(B_{x,\sigma(x)})$  for all  $x \in \Lambda$  with the derivatives measurable in  $x$  and with  $\|\mathcal{F}_x\|_{C^{N,\alpha}}$   $\varepsilon$ -tempered on  $\Lambda$  then

(1) There exists a positive measurable function  $\rho : \Lambda \rightarrow \mathbb{R}$  which is  $\kappa\varepsilon$ -tempered on  $\Lambda$  and a measurable family  $\{\mathcal{H}_x\}_{x \in \Lambda}$  of  $C^{N,\alpha}$  diffeomorphisms  $\mathcal{H}_x : B_{x,\rho(x)} \rightarrow \mathcal{E}_x$  satisfying  $\mathcal{H}_x(0) = 0$  and  $D_0\mathcal{H}_x = \text{Id}$  which conjugate  $\mathcal{F}$  to a sub-resonance polynomial extension  $\mathcal{P}$ :

$$\mathcal{H}_{f_x} \circ \mathcal{F}_x = \mathcal{P}_x \circ \mathcal{H}_x, \quad \text{where } \mathcal{P}_x \in \mathcal{S}_{x,f_x} \text{ for all } x \in \Lambda.$$

Moreover,  $\|\mathcal{H}_x\|_{C^{N,\alpha}(B_{x,\rho(x)})}$  is  $L\varepsilon$ -tempered on  $\Lambda$  and  $\|D_0^{(n)}\mathcal{H}_x\|$  is  $n^2\varepsilon$ -tempered on  $\Lambda$  for  $n = 1, \dots, N$ , with respect to the  $\varepsilon$ -Lyapunov metric (3.2).

(2) Suppose  $\tilde{\mathcal{H}} = \{\tilde{\mathcal{H}}_x\}_{x \in \Lambda}$  is another measurable family of diffeomorphisms as in (1) conjugating  $\mathcal{F}$  to a sub-resonance polynomial extension  $\tilde{\mathcal{P}}$ . Then for all  $x \in \Lambda$  there exists  $G_x \in G_x^\chi$  which is measurable and tempered in  $x$  such that  $\mathcal{H}_x = G_x \circ \tilde{\mathcal{H}}_x$ . Moreover, if  $D_0^{(n)}\tilde{\mathcal{H}}_x = D_0^{(n)}\mathcal{H}_x$  for all  $n = 2, \dots, d = \lfloor \chi_1/\chi_\ell \rfloor$ , then  $\mathcal{H}_x = \tilde{\mathcal{H}}_x$  for all  $x \in \Lambda$ . In particular,  $\{\mathcal{H}_x\}_{x \in \Lambda}$  is unique if  $d = 1$ .

(3) Let  $g : X \rightarrow X$  be an invertible map commuting with  $f$  and let  $\Lambda'$  be a subset of  $\Lambda$  which is both  $f$  and  $g$  invariant. Let  $\mathcal{G}(x, t) = (g(x), \mathcal{G}_x(t))$  be an extension of  $g$  to  $\mathcal{E}$  which preserves the zero section and commutes with  $\mathcal{F}$ . Suppose that  $\mathcal{G}_x$  is  $C^{N, \alpha}(B_{x, \sigma(x)})$  for all  $x \in \Lambda'$  with the derivatives measurable in  $x$ , and that  $\|\mathcal{G}_x\|_{C^{N, \alpha}}$  and  $\|(D_0 \mathcal{G}_x)^{-1}\|$  are  $\varepsilon$ -tempered on  $\Lambda'$ . Then  $\mathcal{H}_{gx} \circ \mathcal{G}_x \circ \mathcal{H}_x^{-1} \in S_{x, fx}$  for all  $x \in \Lambda'$ .

**Corollary 2.4.** Suppose that  $\mathcal{F}_x$  is  $C^\infty(B_{x, \sigma(x)})$  and that  $1/\sigma$  and  $\|\mathcal{F}_x\|_{C^N}$  for each  $n \in \mathbb{N}$  are  $\varepsilon$ -tempered on  $\Lambda$  for each  $\varepsilon > 0$ . Then  $\mathcal{H}_x$  in part (1) of Theorem 2.3 is  $C^\infty(B_{x, \rho(x)})$ .

**Normal forms on stable manifolds.** Let  $\mathcal{M}$  be a smooth manifold and let  $f$  be a diffeomorphism of  $\mathcal{M}$  preserving an ergodic Borel probability measure  $\mu$ . We assume that  $f$  is  $C^{N, \alpha}$ , that is  $C^N$  with  $N^{th}$  derivative  $\alpha$ -Hölder on  $\mathcal{M}$ . We denote by  $\Lambda$  the full measure set of Lyapunov regular points for  $(f, \mu)$ . Let  $\chi_1 < \dots < \chi_\ell$  be the Lyapunov exponents of  $(f, \mu)$  and suppose  $\ell$  is such that  $\chi_\ell < 0$ . Then for each  $x \in \Lambda$  there exists the (strong) stable manifold  $W_x$  tangent to  $\mathcal{E}_x = \mathcal{E}_x^1 \oplus \dots \oplus \mathcal{E}_x^\ell$  [R79, Theorem 6.1].

**Theorem 2.5** (Normal forms on stable manifolds). *Let  $\mathcal{M}$  be a smooth manifold and let  $f$  be a  $C^{N, \alpha}$  diffeomorphism of  $\mathcal{M}$  preserving an ergodic Borel probability measure  $\mu$ . Suppose that  $N \geq 1$ ,  $0 < \alpha \leq 1$  and  $N + \alpha > \chi_1/\chi_\ell$ . Then there exist a full measure set  $X$  which consists of full stable manifolds  $W_x$  and a measurable family  $\{\mathcal{H}_x\}_{x \in X}$  of  $C^{N, \alpha}$  diffeomorphisms*

$$\mathcal{H}_x : W_x \rightarrow \mathcal{E}_x = T_x W_x \quad \text{such that}$$

- (i)  $\mathcal{P}_x = \mathcal{H}_{fx} \circ f \circ \mathcal{H}_x^{-1} : \mathcal{E}_x \rightarrow \mathcal{E}_{fx}$  is a sub-resonance polynomial map for each  $x \in X$ ,
- (ii)  $\mathcal{H}_x(x) = 0$  and  $D_x \mathcal{H}_x$  is the identity map for each  $x \in X$ ,
- (iii)  $\|\mathcal{H}_x\|_{C^{N, \alpha}}$  is tempered on  $X$ ,
- (iv)  $\mathcal{H}_y \circ \mathcal{H}_x^{-1} : \mathcal{E}_x \rightarrow \mathcal{E}_y$  is a sub-resonance polynomial map for all  $x \in X$  and  $y \in W_x$ ,
- (v) If  $g : \mathcal{M} \rightarrow \mathcal{M}$  is a  $C^{N, \alpha}$  diffeomorphism commuting with  $f$  which preserves the measure class of  $\mu$  then  $\mathcal{H}_{gx} \circ \mathcal{G}_x \circ \mathcal{H}_x^{-1} : \mathcal{E}_x \rightarrow \mathcal{E}_{gx}$  is a sub-resonance polynomial map for all  $x$  in a full measure set  $X'$  which consists of full stable manifolds.

Another way to interpret (iv) is to view  $\mathcal{H}_x$  as a coordinate chart on  $W_x$  identifying it with  $\mathcal{E}_x$ . In this coordinate chart, (iv) yields that all transition maps  $\mathcal{H}_y \circ \mathcal{H}_x^{-1}$  for  $y \in W_x$  are in the group  $\bar{G}_x^\chi$  generated by  $G_x^\chi$  and the translations of  $\mathcal{E}_x$ . Thus  $\mathcal{H}_x$  gives the leaf a structure of homogeneous space  $W_x \sim \bar{G}_x^\chi / G_x^\chi$ , which is consistent with other coordinate charts  $\mathcal{H}_y$  for  $y \in W_x$  and is preserved by the normal form  $\mathcal{P}_x$  by (i).

**Corollary 2.6.** *Under the assumptions of the Theorem 2.5, if  $d = \lfloor \chi_1/\chi_\ell \rfloor = 1$ , i.e.  $2\chi_\ell < \chi_1$ , then  $\mathcal{P}_x$  is the linear map  $Df|_{\mathcal{E}_x}$ , the family  $\{\mathcal{H}_x\}_{x \in X}$  satisfying (ii) and (iii) is unique, the maps  $\mathcal{H}_y \circ \mathcal{H}_x^{-1} : \mathcal{E}_x \rightarrow \mathcal{E}_y$  are affine for all  $x \in X$  and  $y \in W_x$ , and  $\mathcal{H}_y$  depends  $C^N$ -smoothly on  $y$  along the stable manifolds.*

## 3. LYAPUNOV EXPONENTS AND LYAPUNOV NORM

In this section we review some basic definitions and facts of the Oseledets theory of linear extensions. We use [BP] as a general reference. For a linear extension  $F$  of a map  $f$  we will use the notation

$$(3.1) \quad F_x^n = F_{f^{n-1}x} \circ \cdots \circ F_{fx} \circ F_x.$$

**Theorem 3.1** (Oseledets Multiplicative Ergodic Theorem, see [BP] Theorem 3.4.3). *Let  $f$  be an invertible ergodic measure-preserving transformation of a Lebesgue probability space  $(X, \mu)$ . Let  $F$  be a measurable linear extension satisfying  $\log \|F_x\| \in L^1(X, \mu)$  and  $\log \|F_x^{-1}\| \in L^1(X, \mu)$ . Then there exist numbers  $\chi_1 < \cdots < \chi_\ell$ , an  $f$ -invariant set  $\Lambda$  with  $\mu(\Lambda) = 1$ , and an  $F$ -invariant Lyapunov decomposition*

$$\mathcal{E}_x = \mathcal{E}_x^1 \oplus \cdots \oplus \mathcal{E}_x^\ell \text{ for } x \in \Lambda$$

such that

- (i)  $\lim_{n \rightarrow \pm\infty} n^{-1} \log \|F_x^n v\| = \chi_i$  for any  $i = 1, \dots, \ell$  and any  $0 \neq v \in \mathcal{E}_x^i$ , and
- (ii)  $\lim_{n \rightarrow \pm\infty} n^{-1} \log \det F_x^n = \sum_{i=1}^\ell m_i \chi_i$ , where  $m_i = \dim \mathcal{E}_x^i$ .

The numbers  $\chi_1, \dots, \chi_\ell$  are called the *Lyapunov exponents* of  $F$  and the points of the set  $\Lambda^\mu$  are called *regular*.

We denote the standard scalar product in  $\mathbb{R}^m$  by  $\langle \cdot, \cdot \rangle$ . For a fixed  $\varepsilon > 0$  and a regular point  $x$ , the  $\varepsilon$ -Lyapunov scalar product (or metric)  $\langle \cdot, \cdot \rangle_{x,\varepsilon}$  in  $\mathcal{E}_x = \mathbb{R}^m$  is defined as follows. For  $u \in \mathcal{E}_x^i$  and  $v \in \mathcal{E}_x^j$  with  $i \neq j$ ,  $\langle u, v \rangle_{x,\varepsilon} := 0$ , and for  $i = 1, \dots, \ell$  and  $u, v \in \mathcal{E}_x^i$ ,

$$(3.2) \quad \langle u, v \rangle_{x,\varepsilon} = m \sum_{n \in \mathbb{Z}} \langle F_x^n(u), F_x^n(v) \rangle \exp(-2\chi_i n - \varepsilon|n|).$$

Note that the series converges exponentially for any regular  $x$ . The constant  $m$  in front of the conventional formula is introduced for more convenient comparison with the standard scalar product. Usually,  $\varepsilon$  will be fixed and we will denote  $\langle \cdot, \cdot \rangle_{x,\varepsilon}$  simply by  $\langle \cdot, \cdot \rangle_x$  and call it the *Lyapunov scalar product*. The norm generated by this scalar product is called the *Lyapunov norm* and is denoted by  $\|\cdot\|_{x,\varepsilon}$  or  $\|\cdot\|_x$ .

Below we summarize the basic properties of the Lyapunov scalar product and norm, for more details see [BP, Sections 3.5.1-3.5.3]. A direct calculation shows [BP, Theorem 3.5.5] that for any regular  $x$  and any  $u \in \mathcal{E}_x^i$

$$(3.3) \quad \exp(n\chi_i - \varepsilon|n|) \|u\|_{x,\varepsilon} \leq \|F_x^n(u)\|_{f^n x, \varepsilon} \leq \exp(n\chi_i + \varepsilon|n|) \|u\|_{x,\varepsilon} \quad \text{for all } n \in \mathbb{Z},$$

$$(3.4) \quad \exp(n\chi_\ell - \varepsilon n) \leq \|F_x^n\|_{f^n x \leftarrow x} \leq \exp(n\chi_\ell + \varepsilon n) \quad \text{for all } n \in \mathbb{N},$$

where  $\|\cdot\|_{f^n x \leftarrow x}$  is the operator norm with respect to the Lyapunov norms. It is defined for any points  $x, y \in \Lambda$  and any linear map  $F : \mathcal{E}_x \rightarrow \mathcal{E}_y$  as follows

$$\|F\|_{y \leftarrow x} = \sup \{ \|Fu\|_{y,\varepsilon} : u \in \mathcal{E}_x, \|u\|_{x,\varepsilon} = 1 \}.$$

We emphasize that Lyapunov scalar product and norm are defined only for regular points and depend measurably on the point. Thus, a comparison with the standard norm is important. The uniform lower bound follows easily from the definition:  $\|u\|_{x,\varepsilon} \geq \|u\|$ . The upper bound is not uniform, but it changes slowly along the regular orbits [BP, Proposition 3.5.8]: there exists a measurable function  $K_\varepsilon(x)$  defined on the set of regular points  $\Lambda$  such that

$$(3.5) \quad \|u\| \leq \|u\|_{x,\varepsilon} \leq K_\varepsilon(x)\|u\| \quad \text{for all } x \in \Lambda \text{ and } u \in \mathcal{E}_x, \quad \text{and}$$

$$(3.6) \quad K_\varepsilon(x)e^{-\varepsilon|n|} \leq K_\varepsilon(f^n x) \leq K_\varepsilon(x)e^{\varepsilon|n|} \quad \text{for all } x \in \Lambda \text{ and } n \in \mathbb{Z}.$$

These estimates are obtained in [BP] using the fact that  $\|u\|_{x,\varepsilon}$  is *tempered*, but they can also be verified directly using the definition of  $\|u\|_{x,\varepsilon}$  on each Lyapunov space and noting that angles between the spaces change slowly.

Using (3.5) we obtain that for any point  $x, y \in \Lambda$  and any linear map  $F : \mathcal{E}_x \rightarrow \mathcal{E}_y$

$$(3.7) \quad K_\varepsilon(x)^{-1}\|F\| \leq \|F\|_{y \leftarrow x} \leq K_\varepsilon(y)\|F\|.$$

When  $\varepsilon$  is fixed we will usually omit it and write  $K(x) = K_\varepsilon(x)$  and  $\|u\|_x = \|u\|_{x,\varepsilon}$ .

Similarly, we will consider the Lyapunov norm of a homogeneous polynomial map  $R : \mathcal{E}_x \rightarrow \mathcal{E}_y$  of order  $n$  defined as

$$(3.8) \quad \|R\|_{y \leftarrow x} = \sup \{ \|R(u)\|_{y,\varepsilon} : u \in \mathcal{E}_x, \|u\|_{x,\varepsilon} = 1 \}.$$

It follows that

$$(3.9) \quad \|R \circ P\| \leq \|R\| \cdot \|P\|^n.$$

For a homogeneous polynomial map  $P : \mathcal{E}_x \rightarrow \mathcal{E}_y$  of order  $n$  we have

$$(3.10) \quad K_\varepsilon(x)^n\|P\| \leq \|P\|_{y \leftarrow x} \leq K_\varepsilon(y)\|P\|.$$

This formula allows us to switch between the standard and Lyapunov norms in spaces of polynomials and smooth functions.

#### 4. PROOF OF THEOREM 2.3

We give the proof for the case  $\alpha > 0$ . The proof for  $\alpha = 0$  with  $N \geq 2$  is similar but avoids difficulties of estimating the Hölder constant at 0.

We denote  $\mathcal{F}_x^n = \mathcal{F}_{f^{n-1}x} \circ \cdots \circ \mathcal{F}_{fx} \circ \mathcal{F}_x$ . We take  $L = \max \{ \kappa, M + N^3 + 2N^2 \}$ , where  $M = M(d)$  is chosen to satisfy (4.29). We set  $\varepsilon_* = \varepsilon_0/3(N+1)$ , where

$$(4.1) \quad \varepsilon_0 = \min \{ \nu/(2L + 4(N+1+\alpha)), -\chi_\ell/(2L+2), -\lambda/(N^2 + N+1) \} > 0, \\ \text{where } \nu = \chi_1 - (N+\alpha)\chi_\ell > 0 \quad \text{and } \lambda < 0 \text{ is given by (4.15).}$$

We fix  $\varepsilon < \varepsilon_0$  and let  $K = K_\varepsilon$  be as in (3.5). Since  $\|\mathcal{F}_x\|_{C^{N,\alpha}}$  is  $\varepsilon$ -tempered, there is a function  $C : \Lambda \rightarrow [1, \infty)$  such that for all  $x \in \Lambda$  and  $n \in \mathbb{N}$

$$(4.2) \quad \|\mathcal{F}_x\|_{C^{N,\alpha}} \leq C(x) \quad \text{and} \quad C(f^n x) \leq e^{n\varepsilon}C(x).$$

Similarly, replacing  $\sigma$  by a smaller function if necessary, we can assume that it satisfies

$$(4.3) \quad \sigma : \Lambda \rightarrow (0, 1] \quad \text{and} \quad \sigma(f^n x) \geq e^{-n\varepsilon} \sigma(x).$$

**Lemma 4.1.** *Under the assumptions of Theorem 2.3, there exists a function  $\rho : \Lambda \rightarrow \mathbb{R}_+$  such that for all  $x \in \Lambda$ ,  $n \in \mathbb{N}$ , and  $t \in B_{x, \rho(x)} \subset \mathcal{E}_x$ , we have  $\rho(x) < \sigma(x) \leq 1$  and*

$$(1) \quad \rho(f^n x) \geq e^{-\kappa \varepsilon n} \rho(x), \text{ where } \kappa \text{ is given by (2.7),}$$

$$(2) \quad \|D_t \mathcal{F}_x^n\|_{f^n x \leftarrow x} \leq e^{(\chi_\ell + 2\varepsilon)n},$$

$$(3) \quad \|D_t \mathcal{F}_x^n\| \leq K(x) e^{(\chi_\ell + 2\varepsilon)n},$$

$$(4) \quad \|\mathcal{F}_x^n(t)\| \leq K(x) e^{(\chi_\ell + 2\varepsilon)n} \|t\|,$$

$$(5) \quad \|\mathcal{F}_x^n(t)\|_{f^n x} \leq e^{(\chi_\ell + 2\varepsilon)n} \|t\|_x.$$

*Proof.* We take  $\beta = 1$  if  $N \geq 2$  and  $\beta = \alpha > 0$  if  $N = 1$ . For each  $x \in \Lambda$  we define

$$(4.4) \quad \rho(x) = \sigma(x) [\varepsilon e^{\chi_\ell} (C(x) K(x)^2)^{-1}]^{1/\beta}.$$

Then (1) follows from (4.2), (4.3), and (3.6); (5) follows from (2). We prove (2), (3), and (4) by induction. The statements are clear for  $n = 0$ , suppose they hold for  $n$ . Note that (2) implies (3) by (3.7), and (3) implies (4). We observe that

$$\|D_t \mathcal{F}_x^{n+1}\|_{f^{n+1}x \leftarrow x} \leq \|D_{t'} \mathcal{F}_{f^n x}\|_{f^{n+1}x \leftarrow f^n x} \cdot \|D_t \mathcal{F}_{f^n x}\|_{f^n x \leftarrow x}, \quad \text{where } t' = \mathcal{F}_x^n(t).$$

Then (2) follows from the inductive assumption and

$$(4.5) \quad \|D_t \mathcal{F}_{f^n x}\|_{f^{n+1}x \leftarrow f^n x} \leq e^{\chi_\ell + 2\varepsilon}.$$

To prove (4.5) we denote  $\Delta = D_{t'} \mathcal{F}_{f^n x} - D_0 \mathcal{F}_{f^n x}$ . By the choice of  $\beta$ , the  $\beta$ -Hölder constant of  $D_s \mathcal{F}_{f^n x}$  at 0 is at most  $C(f^n x)$ , so using (3.6) we obtain

$$\|\Delta\|_{f^{n+1}x \leftarrow f^n x} \leq K(f^{n+1}x) \|\Delta\| \leq K(f^{n+1}x) C(f^n x) \|t'\|^\beta \leq$$

and using (4.2) and the inductive assumption (4) we get

$$\leq K(x) C(x) e^{(2n+1)\varepsilon} K(x) e^{\beta(\chi_\ell + 2\varepsilon)n} \|t\|^\beta \leq e^\varepsilon C(x) K(x)^2 e^{[2\varepsilon + \beta(\chi_\ell + 2\varepsilon)]n} \|t\|^\beta.$$

Since  $\|t\| \leq \rho(x)$  and  $\beta\chi_\ell + 2(1 + \beta)\varepsilon \leq 0$  we obtain

$$\|\Delta\|_{f^{n+1}x \leftarrow f^n x} \leq e^\varepsilon C(x) K(x)^2 \rho(x)^\beta \leq \varepsilon e^{\chi_\ell + \varepsilon} \sigma(x)^\beta \leq \varepsilon e^{\chi_\ell + \varepsilon}.$$

Since

$$D_0 \mathcal{F}_{f^n x} = F_{f^n x} \quad \text{and} \quad \|F_{f^n x}\|_{f^{n+1}x \leftarrow f^n x} \leq e^{\chi_\ell + \varepsilon}$$

by (3.4), we conclude that

$$\|D_{t'} \mathcal{F}_{f^n x}\|_{f^{n+1}x \leftarrow f^n x} \leq \|\Delta\|_{f^{n+1}x \leftarrow f^n x} + \|F_{f^n x}\|_{f^{n+1}x \leftarrow f^n x} \leq \varepsilon e^{\chi_\ell + \varepsilon} + e^{\chi_\ell + \varepsilon} \leq e^{\chi_\ell + 2\varepsilon}.$$

□



#### 4.1. Construction of $\mathcal{P}$ and of the Taylor polynomial for $\mathcal{H}$ .

For each  $x \in \Lambda$  and map  $\mathcal{F}_x : \mathcal{E}_x \rightarrow \mathcal{E}_{f_x}$  we consider the Taylor polynomial at  $t = 0$ :

$$(4.6) \quad \mathcal{F}_x(t) \sim \sum_{n=1}^N F_x^{(n)}(t).$$

As a function of  $t$ ,  $F_x^{(n)}(t) : \mathcal{E}_x \rightarrow \mathcal{E}_{f_x}$  is a homogeneous polynomial map of degree  $n$ . First we construct the Taylor polynomials at  $t = 0$  for the desired coordinate change  $\mathcal{H}_x(t)$  and the polynomial extension  $\mathcal{P}_x(t)$ . We use similar notations for these Taylor polynomials:

$$\mathcal{H}_x(t) \sim \sum_{n=1}^N H_x^{(n)}(t) \quad \text{and} \quad \mathcal{P}_x(t) = \sum_{n=1}^d P_x^{(n)}(t).$$

For the first derivative we choose

$$H_x^{(1)} = \text{Id} : \mathcal{E}_x \rightarrow \mathcal{E}_x \quad \text{and} \quad P_x^{(1)} = F_x \quad \text{for all } x \in \Lambda.$$

We construct the terms  $H_x^{(n)}$  inductively to ensure that the terms  $P_x^{(n)}$  determined by the conjugacy equation are of sub-resonance type. The base of the induction is the linear terms chosen above. For each  $x \in \Lambda$  we will construct  $H_x^{(n)}$  and  $P_x^{(n)}$  which are measurable in  $x$  and  $n^2\varepsilon$ -tempered, i.e.

$$(4.7) \quad \sup_{k \in \mathbb{N}} e^{-n^2\varepsilon k} \|H_{f^k x}^{(n)}\|_{f^k x \leftarrow f^k x} < \infty \quad \text{and} \quad \sup_{k \in \mathbb{N}} e^{-n^2\varepsilon k} \|P_{f^k x}^{(n)}\|_{f^k x \leftarrow f^k x} < \infty.$$

Using these notations in the conjugacy equation  $\mathcal{H}_{f_x} \circ \mathcal{F}_x = \mathcal{P}_x \circ \mathcal{H}_x$

$$\left( \text{Id} + \sum_{i=2}^N H_{f_x}^{(i)} \right) \circ \left( F_x + \sum_{i=2}^N F_x^{(i)} \right) \sim \left( F_x + \sum_{i=2}^d P_x^{(i)} \right) \circ \left( \text{Id} + \sum_{i=2}^N H_x^{(i)} \right).$$

and considering the terms of degree  $N \geq n \geq 2$ , we obtain

$$F_x^{(n)} + H_{f_x}^{(n)} \circ F(x) + \sum H_{f_x}^{(i)} \circ F_x^{(j)} = F_x \circ H_x^{(n)} + P_x^{(n)} + \sum P_x^{(j)} \circ H_x^{(i)},$$

where the summations are over all  $i$  and  $j$  such that  $ij = n$  and  $1 < i, j < n$ . We rewrite the equation as

$$(4.8) \quad F_x^{-1} \circ P_x^{(n)} = -H_x^{(n)} + F_x^{-1} \circ H_{f_x}^{(n)} \circ F_x + Q_x,$$

where

$$(4.9) \quad Q_x = F_x^{-1} \left( F_x^{(n)} + \sum_{ij=n, 1 < i, j < n} H_{f_x}^{(i)} \circ F_x^{(j)} - P_x^{(j)} \circ H_x^{(i)} \right).$$

We note that  $Q_x$  is composed only of terms  $H_x^{(i)}$  and  $P_x^{(i)}$  with  $1 < i < n$ , which are already constructed, and terms  $F_x^{(i)}$  with  $1 < i \leq n$ , which are given. Thus by the inductive assumption  $Q_x$  is defined for all  $x \in \Lambda$  and is measurable and tempered in  $x$ .

Let  $\mathcal{R}_x^{(n)}$  be the space of all polynomial maps on  $\mathcal{E}_x$  of degree  $n$ , and let  $\mathcal{S}_x^{(n)}$  and  $\mathcal{N}_x^{(n)}$  be the subspaces of sub-resonance and non sub-resonance polynomials respectively. We seek  $H_x^{(n)}$  so that the right side of (4.8) is in  $\mathcal{S}_x^{(n)}$ , and hence so is  $P_x^{(n)}$  when defined by this equation.

Projecting (4.8) to the factor bundle  $\mathcal{R}^{(n)}/\mathcal{S}^{(n)}$ , our goal is to solve the equation

$$(4.10) \quad 0 = -\bar{H}_x^{(n)} + F_x^{-1} \circ \bar{H}_{fx}^{(n)} \circ F_x + \bar{Q}_x,$$

where  $\bar{H}^{(n)}$  and  $\bar{Q}$  are the projections of  $H^{(n)}$  and  $Q$  respectively.

We consider the bundle automorphism  $\Phi : \mathcal{R}^{(n)} \rightarrow \mathcal{R}^{(n)}$  covering  $f^{-1} : \mathcal{M} \rightarrow \mathcal{M}$  given by the maps  $\Phi_x : \mathcal{R}_{fx}^{(n)} \rightarrow \mathcal{R}_x^{(n)}$

$$(4.11) \quad \Phi_x(R) = F_x^{-1} \circ R \circ F_x.$$

Since  $F$  preserves the splitting  $\mathcal{E} = \mathcal{E}^1 \oplus \dots \oplus \mathcal{E}^\ell$ , it follows from the definition that the sub-bundles  $\mathcal{S}^{(n)}$  and  $\mathcal{N}^{(n)}$  are  $\Phi$ -invariant. We denote by  $\tilde{\Phi}$  the induced automorphism of  $\mathcal{R}^{(n)}/\mathcal{S}^{(n)}$  and conclude that (4.10) is equivalent to

$$(4.12) \quad \bar{H}_x^{(n)} = \tilde{\Phi}_x(\bar{H}_{fx}^{(n)}), \quad \text{where } \tilde{\Phi}_x(R) = \bar{\Phi}_x(R) + \bar{Q}_x.$$

Thus a solution of (4.10) is a  $\tilde{\Phi}$ -invariant section of  $\mathcal{R}^{(n)}/\mathcal{S}^{(n)}$ . We will show that  $\tilde{\Phi}$  is a nonuniform contraction and that it has a unique measurable tempered invariant section. First, for polynomials of specific homogeneous type the exponent of  $\Phi$  is determined by the exponents of  $F$  as follows.

**Lemma 4.2.** *For a polynomial  $R : \mathcal{E}_{fx} \rightarrow \mathcal{E}_{fx}^i$  of homogeneous type  $s = (s_1, \dots, s_\ell)$*

$$(4.13) \quad \|\Phi_x(R)\|_{x \leftarrow x} \leq e^{-\chi_i + \sum s_j \chi_j + (n+1)\varepsilon} \|R\|_{fx \leftarrow fx}.$$

*Proof.* Suppose that  $v = v_1 + \dots + v_\ell$ , where  $v_j \in \mathcal{E}_x^j$ , and  $\|v\|_x = 1$ . We denote  $a_j = \|F|_{\mathcal{E}_x^j}\|_{fx \leftarrow x}$  and observe that  $F_x(v_j) = a_j v'_j \in \mathcal{E}_{fx}^j$  with  $\|v'_j\|_{fx} \leq \|v_j\|_x$ . Since  $R$  has homogeneous type  $s = (s_1, \dots, s_\ell)$  we obtain by (2.1) that

$$(4.14) \quad (R \circ F_x)(v) = R(a_1 v'_1 + \dots + a_\ell v'_\ell) = a_1^{s_1} \dots a_\ell^{s_\ell} \cdot R(v'_1 + \dots + v'_\ell).$$

where  $v' = v'_1 + \dots + v'_\ell$  has  $\|v'\|_{fx} \leq \|v\|_x = 1$  by orthogonality of the splitting in the Lyapunov metric. Thus

$$\|(R \circ F_x)(v)\|_{fx} = a_1^{s_1} \dots a_\ell^{s_\ell} \cdot \|R(v')\|_{fx} \leq a_1^{s_1} \dots a_\ell^{s_\ell} \cdot \|R\|_{fx \leftarrow fx}$$

for any  $v \in \mathcal{E}_x$  with  $\|v\|_x = 1$ , so we get  $\|R \circ F_x\|_{x \leftarrow fx} \leq a_1^{s_1} \dots a_\ell^{s_\ell} \|R\|_{fx \leftarrow fx}$  by definition (3.8). Now (3.9) yields

$$\begin{aligned} \|\Phi_x(R)\|_{x \leftarrow x} &= \|F|_{\mathcal{E}_x^i}^{-1} \circ R \circ F_x\|_{x \leftarrow x} \leq \|F|_{\mathcal{E}_x^i}^{-1}\|_{x \leftarrow fx} \cdot \|R \circ F_x\|_{x \leftarrow fx} \leq \\ &\leq \|F|_{\mathcal{E}_x^i}^{-1}\|_{x \leftarrow fx} \cdot a_1^{s_1} \dots a_\ell^{s_\ell} \cdot \|R\|_{fx \leftarrow fx} \leq e^{-\chi_i + \varepsilon} \cdot \prod_j (e^{\chi_j + \varepsilon})^{s_j} \cdot \|R\|_{fx \leftarrow fx}. \end{aligned}$$

Since  $a_j = \|F|_{\mathcal{E}_x^j}\|_{fx \leftarrow x} \leq e^{\chi_j + \varepsilon}$  and  $\|F|_{\mathcal{E}_x^i}^{-1}\|_{x \leftarrow fx} \leq e^{-\chi_i + \varepsilon}$  by (3.3).  $\square$

**Remark.** Similarly, one can show that  $\|\Phi^{-1}(R)\| \leq e^{\chi_i - \sum s_j \chi_j + (n+1)\varepsilon}$ . Since this holds for any  $\varepsilon > 0$ , one can obtain that the Lyapunov exponent of  $\Phi$  on  $R$  is

$$\lim_{n \rightarrow \pm\infty} n^{-1} \log \|\Phi^n(R)\| = -\chi_i + \sum s_j \chi_j.$$

Taking the supremum of  $-\chi_i + \sum s_j \chi_j$  over all *non* sub-resonance types  $(i; s_1, \dots, s_\ell)$ , that is those for which this value is negative, we define

$$(4.15) \quad \lambda = \sup \left\{ -\chi_i + \sum s_j \chi_j \right\} < 0.$$

We note that  $\lambda < 0$  since there only finitely many such values which are greater than a given number. Thus we obtain the following lemma.

**Lemma 4.3.** *The map  $\Phi : \mathcal{N}^{(n)} \rightarrow \mathcal{N}^{(n)}$  given by (4.11) is a nonuniform contraction over  $f^{-1}$ , and hence so is  $\tilde{\Phi} : \mathcal{R}^{(n)}/\mathcal{S}^{(n)} \rightarrow \mathcal{R}^{(n)}/\mathcal{S}^{(n)}$  given by (4.12). More precisely,  $\|\Phi_x(R)\|_x \leq e^{\lambda + (n+1)\varepsilon} \|R\|_{fx}$ .*

*Proof.* The statement about  $\tilde{\Phi}$  follows since the linear part  $\bar{\Phi}$  of  $\tilde{\Phi}$  is given by  $\Phi$  when  $\mathcal{R}^{(n)}/\mathcal{S}^{(n)}$  is naturally identified with  $\mathcal{N}^{(n)}$ .  $\square$

It follows from the previous remark that  $\lambda$  is the maximal Lyapunov exponent of  $\Phi$  over  $f^{-1}$  on the space of non sub-resonant polynomials, and that all Lyapunov exponents of  $\Phi|_{\mathcal{S}^{(n)}}$  are non-negative.

Now we construct a  $\tilde{\Phi}$ -invariant measurable section of the bundle  $\mathcal{B} = \mathcal{R}^{(n)}/\mathcal{S}^{(n)}$  and study its properties. The construction is orbit-wise. We fix a point  $x \in \Lambda$ , consider its positive orbit  $\{x_k = f^k x : k \geq 0\}$ , and define the Banach space

$$\mathcal{B}^x = \{R = (R_k)_{k=0}^\infty : R_k \in \mathcal{B}_{x_k}, \|R\| < \infty\}, \quad \text{where } \|R\| = \sup_{k \geq 0} e^{-\varepsilon n^2 k} \|R_k\|_{x_k \leftarrow x_k}$$

and  $\|R_k\|_{x_k \leftarrow x_k}$  is the norm induced on  $\mathcal{B}_{x_k}$  by the Lyapunov norm  $\|\cdot\|_{x_k}$  on  $\mathcal{E}_x$ . We denote  $\tilde{Q} = (\tilde{Q}_{x_k})_{k=0}^\infty$  and claim that it is in  $\mathcal{B}^x$ . For this we need to estimate how the Lyapunov norm of (4.9) grows along the trajectory:

$$(4.16) \quad \begin{aligned} \|Q_{x_k}\|_{x_k \leftarrow x_k} &\leq \|F_{x_k}^{-1}\|_{x_k \leftarrow x_{k+1}} \cdot (\|F_{x_k}^{(n)}\|_{x_{k+1} \leftarrow x_k} + \\ &\sum_{ij=n, 2 \leq i, j \leq n/2} \|H_{x_{k+1}}^{(i)}\|_{x_{k+1} \leftarrow x_{k+1}} \|F_{x_k}^{(j)}\|_{x_{k+1} \leftarrow x_k}^i + \|P_{x_k}^{(j)}\|_{x_{k+1} \leftarrow x_k} \|H_{x_k}^{(i)}\|_{x_k \leftarrow x_k}^j). \end{aligned}$$

First  $\|F_{x_k}^{-1}\|_{x_k \leftarrow x_{k+1}} \leq e^{-\chi_\ell + \varepsilon}$  for all  $x$  and  $k$  by (3.4). The exponential growth rate in  $k$  of  $\|F_{x_k}^{(n)}\|_{x_{k+1} \leftarrow x_k}$  is at most  $2\varepsilon$ . Indeed, using (3.10) and (3.6) we can obtain from (4.2) the corresponding estimate for  $C^{N,\alpha}$  norm with respect to the Lyapunov metric on  $\mathcal{E}_{x_k}$ :

$$(4.17) \quad \|\mathcal{F}_{x_k}\|_{C^{N,\alpha}, x_k} \leq K(fx_k) \|\mathcal{F}_{x_k}\|_{C^{N,\alpha}} \leq K(x_{k+1}) C(x_k) \leq e^{(2k+1)\varepsilon} K(x) C(x).$$

Then using the inductive assumption (4.7) we can estimate the exponential growth rate of the two terms in the sum respectively as  $(i^2 + 2i)\varepsilon$  and  $(j^2 + i^2 j)\varepsilon$ , which are

at most  $((n/2)^2 + in)\varepsilon < n^2\varepsilon$ . So the exponential growth rate of  $\|Q_{x_k}\|_{x_k \leftarrow x_k}$  can be estimated by  $n^2\varepsilon$  and thus  $\|\tilde{Q}\| < \infty$ .

Then  $\tilde{\Phi}^x$  induces an operator on  $\mathcal{B}^x$  by  $(\tilde{\Phi}^x(R))_k = \bar{\Phi}_{x_k}(R_{k+1}) + \tilde{Q}_k$  and we have

$$\begin{aligned} \|\tilde{\Phi}^x(R) - \tilde{\Phi}^x(R')\| &= \sup_{k \geq 0} e^{-\varepsilon n^2 k} \|\bar{\Phi}_{x_k}(R_{k+1} - R'_{k+1})\|_{x_k \leftarrow x_k} \leq \\ &\leq \sup_{k \geq 0} e^{-\varepsilon n^2 k} e^{\lambda + (n+1)\varepsilon} \|R_{k+1} - R'_{k+1}\|_{x_{k+1} \leftarrow x_{k+1}} \leq \\ &\leq e^{\lambda + (n^2 + n + 1)\varepsilon} \sup_{k \geq 0} e^{-\varepsilon n^2 (k+1)} \|(R_{k+1} - R'_{k+1})\|_{x_{k+1} \leftarrow x_{k+1}} \leq e^{\lambda + (n^2 + n + 1)\varepsilon} \|R - R'\|. \end{aligned}$$

Since  $\lambda + (n^2 + n + 1)\varepsilon < 0$  by the choice of  $\varepsilon$  (4.1),  $\tilde{\Phi}^x$  is a contraction and thus has a unique fixed point  $R^x \in \mathcal{B}^x$ . We claim that  $\bar{H}_x^{(n)} = R_0^x$  is a measurable function which is a unique solution of (4.12) or equivalently (4.10). Measurability follows from the fact that the fixed point can be explicitly written as a series

$$(4.18) \quad \bar{H}_x^{(n)} = \sum_{k=0}^{\infty} (F_x^k)^{-1} \circ \bar{Q}_{x_k} \circ F_x^k.$$

Invariance is clear since  $(R_{k+1}^x)_{k=0}^{\infty}$  is a fixed point of  $\tilde{\Phi}^{fx}$  which coincides with  $(R_k^{fx})_{k=0}^{\infty}$  by uniqueness and thus  $R_1^x = R_0^{fx}$ . More generally,  $\bar{H}_{x_k}^{(n)} = R_0^{x_k} = R_k^x$ , and since  $R^x \in \mathcal{B}^x$ , the exponential growth rate of  $\|\bar{H}_{x_k}^{(n)}\|_{x_k \leftarrow x_k}$  is at most  $n^2\varepsilon$ . Now we can choose  $H_x^{(n)}$  as a lift of  $\bar{H}_x^{(n)}$  to  $\mathcal{R}_x^{(n)}$  which is measurable in  $x$  and satisfies (4.7). Then we define  $P_x^{(n)}$  by equation (4.8). It also satisfies (4.7) as so do  $H$  and  $Q$  and as  $\|F_x\|_{x \leftarrow fx}$  and  $\|F_x^{-1}\|_{fx \leftarrow x}$  are uniformly bounded. This completes the inductive step.

Thus we have constructed the Taylor polynomial for the coordinate change

$$(4.19) \quad \mathcal{H}_x^N(t) = \sum_{n=1}^N H_x^{(n)}(t) \quad \text{of order } N \geq d = \lfloor \chi_1 / \chi_\ell \rfloor$$

and the polynomial map  $\mathcal{P}_x(t) = \sum_{n=1}^d P_x^{(n)}(t)$ .

#### 4.2. Construction of the coordinate change $\mathcal{H}$ .

We rewrite the conjugacy equation  $\mathcal{H}_{fx} \circ \mathcal{F}_x = \mathcal{P}_x \circ \mathcal{H}_x$  in the form

$$(4.20) \quad \mathcal{H}_x = \mathcal{P}_x^{-1} \circ \mathcal{H}_{fx} \circ \mathcal{F}_x.$$

A solution  $\mathcal{H} = \{\mathcal{H}_x\}$  of this equation is a fixed point of the operator  $T$

$$(4.21) \quad T(\mathcal{H})_x = \mathcal{P}_x^{-1} \circ \mathcal{H}_{fx} \circ \mathcal{F}_x.$$

We denote  $R = \mathcal{H} - \mathcal{H}^N$  and observe that  $T(\mathcal{H}) = \mathcal{H}$  if and only if

$$(4.22) \quad T(R) = R + \Delta^N, \quad \text{where } \Delta^N = \mathcal{H}^N - T(\mathcal{H}^N).$$

We will find  $R$  as the fixed point of the contraction

$$(4.23) \quad \tilde{T}(R) = T(R) + \Delta^N$$

in an appropriate space of sequences of functions along an orbit. By the construction of  $\mathcal{H}^N$  and  $\mathcal{P}$ ,  $\mathcal{H}^N$  and  $T(\mathcal{H}^N)$  have the same derivatives at the zero section up to order  $N$ . Hence  $\Delta^N$  has vanishing derivatives at the zero section up to order  $N$ . To define the space we introduce some notations. For any  $x \in \Lambda$  we denote by  $B_{x,r}$  the ball centered at 0 in  $\mathcal{E}_x$  of radius  $r < \rho(x) < 1$  in the Lyapunov norm  $\|\cdot\|_x$ . We define

$$\mathcal{C}_{x,r} = \{R \in C^{N,\alpha}(B_{x,r}, \mathcal{E}_x) : D_0^{(k)}R = 0, k = 0, \dots, N\}.$$

In this proof we will consider the  $C^{N,\alpha}$  norms *with respect to the Lyapunov metric* on  $\mathcal{E}_x$ . They are estimated through the norms for the standard metric (2.5) in (4.17). In particular, we use the  $\alpha$ -Hölder constant (2.4) of  $D^{(N)}R$  at 0 with respect to the Lyapunov metric, which for any  $R \in \mathcal{C}_{x,r}$  is given by

$$(4.24) \quad \|D^{(N)}R\|_{x,\alpha} = \sup\{\|D_t^{(N)}R\|_{x \leftarrow x} \cdot \|t\|_x^{-\alpha} : 0 \neq t \in B_{x,r}\}.$$

Also, for any  $R \in \mathcal{C}_{x,r}$  we can estimate lower derivatives as

$$(4.25) \quad \|D_t^{(m)}R\|_{x \leftarrow x} \leq \|t\|_x^{N-m} \sup\{\|D_s^{(N)}R\|_{x \leftarrow x} : \|s\|_x \leq \|t\|_x\},$$

so using the above Hölder constant we have that for any  $0 \leq m < N$  and  $t \in B_{x,r}$

$$(4.26) \quad \|D_t^{(m)}R\|_{x \leftarrow x} \leq \|t\|_x^{1+\alpha} \|D^{(N)}R\|_{x,\alpha}.$$

Thus the norms of all derivatives are dominated by the Hölder constant and hence

$$(4.27) \quad \|R\|_{C^{N,\alpha}(B_{x,r})} = \|D^{(N)}R\|_{x,\alpha}.$$

So we will take  $\|D^{(N)}R\|_{x,\alpha}$  as the norm  $\mathcal{C}_{x,r}$ , which makes it into a Banach space.

Now we consider the corresponding spaces along the orbit  $x_k = f^k x$ . We will specify later a large  $L > 1$  and a small  $r = r(x) < \rho(x)$ . Using them we define  $r_k = r e^{-2Lk\varepsilon}$  and consider the Banach space

$$\mathcal{C}^x = \{\bar{R} = (R_k)_{k=0}^\infty : R_k \in \mathcal{C}_{x_k, r_k}, \|\bar{R}\|_{\mathcal{C}^x} < \infty\}, \quad \text{where } \|\bar{R}\|_{\mathcal{C}^x} = \sup_{k \geq 0} e^{-Lk\varepsilon} \|D^{(N)}R_k\|_{x_k, \alpha}$$

and the norm  $\|\cdot\|_{x_k, \alpha}$  is defined as in (4.24) and satisfies (4.27). To ensure that  $\bar{\Delta}^N = (\Delta_{x_k}^N)$  is in  $\mathcal{C}^x$  we need to estimate the growth of  $C^{N,\alpha}$  norms of  $\mathcal{H}_x^N$  and  $T(\mathcal{H}_x^N)_x = \mathcal{P}_x^{-1} \circ \mathcal{H}_{fx}^N \circ \mathcal{F}_x$  along the orbit.

We recall that by the construction  $D_0^{(1)}(\mathcal{H}_{x_k}) = \text{Id}$ , and  $D_0^{(1)}(\mathcal{P}_{x_k}) = P_{x_k}^{(1)} = F_{x_k}$ , which satisfy  $\|F_{x_k}\|_{x_{k+1} \leftarrow x_k} \leq e^{\chi_\ell + \varepsilon}$  and  $\|F_{x_k}^{-1}\|_{x_k \leftarrow x_{k+1}} \leq e^{\chi_1 + \varepsilon}$ . Also, for  $2 \leq n \leq d$ , Lyapunov norms of  $D_0^{(n)}(\mathcal{P}_{x_k}) = P_{x_k}^{(n)}$  and  $D_0^{(n)}(\mathcal{H}_{x_k}) = H_{x_k}^{(n)}$  grow at most at the exponential rate  $n^2\varepsilon$  in  $k$  by (4.7).

For  $\mathcal{H}_x^N$ , the derivative of order  $N$  is constant  $H_x^{(N)}$  on  $\mathcal{E}_x$ , and the lower derivatives on  $B_{x, \rho(x)}$  can be inductively estimated by integration similarly to (4.25)

$$\|D_t^{(N-1)}\mathcal{H}_x\|_{x \leftarrow x} \leq \|D_0^{(N-1)}\mathcal{H}_x\|_{x \leftarrow x} + \|t\|_x \|H_x^{(N)}\|_{x \leftarrow x} \leq \|H_x^{(N-1)}\|_{x \leftarrow x} + \|H_x^{(N)}\|_{x \leftarrow x}$$

yielding the same estimate of the exponential rate as for  $H_x^{(N)}$

$$(4.28) \quad \|\mathcal{H}_{x_k}\|_{C^{N,\alpha}(B_{x_k,\rho(x_k)})} \leq c_1(x) e^{N^2 k \varepsilon}.$$

Since  $\mathcal{P}_{x_k}^{-1}$  is also a sub-resonance polynomial, its coefficients can be obtained inductively from those of  $\mathcal{P}_{x_k}$  and hence there exists a constant  $M = M(d) > d$  depending on  $d$  only so that they grow at most at the exponential rate  $M\varepsilon$  in  $k$ . The derivative of order  $d$  is constant on  $\mathcal{E}_{x_k}$ , higher derivatives are zero, and the lower derivatives can be estimated as for  $\mathcal{H}$ , so we obtain for all  $k \geq 0$

$$(4.29) \quad \|(\mathcal{P}_{x_k})^{-1}\|_{C^{N,\alpha}(B_{x_k,\rho(x_k)})} \leq c_2(x) e^{Mk\varepsilon}.$$

To obtain estimates for  $(T(\mathcal{H}^N))_x = \mathcal{P}_x^{-1} \circ \mathcal{H}_x^N \circ \mathcal{F}_x$  we use the following lemma.

**Lemma 4.4.** *If  $Q$  is a polynomial of degree at most  $N$  and  $\mathcal{F}$  is  $C^{N,\alpha}$  then  $Q \circ \mathcal{F}$  is  $C^{N,\alpha}$  and  $\|Q \circ \mathcal{F}\|_{C^{N,\alpha}} \leq c_N \|Q\|_{C^N} \|\mathcal{F}\|_{C^{N,\alpha}}^N$  where  $c_N$  depends on  $N$  only.*

*Proof.* Since  $Q$  is  $C^\infty$  it is clear that  $Q \circ \mathcal{F}$  is  $C^N$ . For the  $N^{\text{th}}$  derivative we have

$$D_t^{(N)}(Q \circ \mathcal{F}) = D_{\mathcal{F}(t)} Q \circ D_t^{(N)} \mathcal{F} + \sum_{kj=N, j < N} D_{\mathcal{F}(t)}^{(k)} Q \circ D_t^{(j)} \mathcal{F}.$$

First we estimate  $\alpha$ -Hölder constant at 0 of the first term. As  $DQ$  is linear, we get

$$D_{\mathcal{F}(t)} Q \circ D_t^{(N)} \mathcal{F} - D_0 Q \circ D_0^{(N)} \mathcal{F} = (D_{\mathcal{F}(t)} Q - D_0 Q) \circ D_t^{(N)} \mathcal{F} + D_0 Q \circ (D_t^{(N)} \mathcal{F} - D_0^{(N)} \mathcal{F})$$

whose norm can be estimated by

$$\begin{aligned} & \|D_{\mathcal{F}(t)} Q - D_0 Q\| \cdot \|D_t^{(N)} \mathcal{F}\| + \|D_0 Q\| \cdot \|D_t^{(N)} \mathcal{F} - D_0^{(N)} \mathcal{F}\| \leq \\ & \|Q\|_{C^2} \cdot \|\mathcal{F}(t)\| \cdot \|\mathcal{F}\|_{C^{N,\alpha}} + \|Q\|_{C^1} \cdot \|\mathcal{F}\|_{C^{N,\alpha}} \cdot \|t\|^\alpha \leq \\ & \|Q\|_{C^2} \cdot \|\mathcal{F}\|_{C^{N,\alpha}} \cdot \|\mathcal{F}\|_{C^1} \cdot \|t\| + \|Q\|_{C^1} \cdot \|\mathcal{F}\|_{C^{N,\alpha}} \cdot \|t\|^\alpha. \end{aligned}$$

So the  $\alpha$ -Hölder constant at 0 of  $D_{\mathcal{F}(t)} Q \circ D_t^{(N)} \mathcal{F}$  is estimated by  $2\|Q\|_{C^N} \|\mathcal{F}\|_{C^{N,\alpha}}^2$ . The other terms in the sum are  $C^1$  and hence are Lipschitz with constant bounded by supremum norms of their derivatives. These norms, along with the norms of lower derivatives of  $Q \circ \mathcal{F}$  can be estimated as a sum of terms of the type

$$\|D_{\mathcal{F}(t)}^{(k)} Q \circ D_t^{(j)} \mathcal{F}\| \leq \|D_{\mathcal{F}(t)}^{(k)} Q\| \cdot \|D_t^{(j)} \mathcal{F}\|^k \leq \|Q\|_{C^N} \|\mathcal{F}\|_{C^{N,\alpha}}^N.$$

We conclude that  $\|Q \circ \mathcal{F}\|_{C^{N,\alpha}} \leq c_N \|Q\|_{C^N} \|\mathcal{F}\|_{C^{N,\alpha}}^N$ .  $\square$

We apply the lemma with  $Q = \mathcal{H}^N$  and then with  $Q = \mathcal{P}_x^{-1}$ . We conclude that  $T(\mathcal{H}^N)$  is  $C^{N,\alpha}$ . Moreover, since  $\|\mathcal{F}\|_{C^{N,\alpha}}^N$  is  $2\varepsilon$ -tempered by (4.17) and using (4.29) and (4.28) we obtain that for  $M' = M + N^3 + 2N^2$

$$(4.30) \quad \|T(\mathcal{H}^N)\|_{C^{N,\alpha}(B_{x_k,\rho(x_k)})} \leq c_3(x) e^{M'k\varepsilon}.$$

Recall that we chose  $L \geq \max\{\kappa, M'\}$  and  $r < \rho(x)$ . Then we obtain by the definition of  $r_k$  and Lemma 4.1 (1) that for all  $k \geq 0$

$$(4.31) \quad r_k = r e^{-2L\varepsilon k} < e^{-L\varepsilon k} \rho(x_k) \leq \rho(x_k).$$

We conclude that with such choices we have  $\bar{\Delta}^N \in \mathcal{C}^x$  with

$$\|\bar{\Delta}^N\|_{\mathcal{C}^x} \leq D' = \sup_{k \geq 0} e^{-Lk\varepsilon} \|\Delta_k\|_{C^{N,\alpha}(B_{x_k, \rho(x_k)})} < \infty.$$

Now we consider the operator induced by  $T$  on  $\mathcal{C}^x$ :

$$(T^x(\bar{R}))_k = (\mathcal{P}_{x_k})^{-1} \circ R_{k+1} \circ \mathcal{F}_{x_k}.$$

We denote by  $B^x(D)$  the ball of radius  $D = D'/\theta$  in  $\mathcal{C}^x$ , where  $\theta > 0$  given by (4.42). We will choose  $L$  and  $r$  so that for any  $\bar{R} \in B^x(D)$  the maps  $(T(R))_k$  are defined on  $B_{x_k, r_k}$  and  $\|T^x(\bar{R})\|_{\mathcal{C}^x} \leq (1 - \theta)\|\bar{R}\|_{\mathcal{C}^x}$ . Then it will follow that  $\tilde{T}^x : B^x(D) \rightarrow B^x(D)$  and is also a contraction, whose unique fixed point gives the desired solution.

First we check that the compositions in  $(T(R))_k$  are well-defined. We take  $t \in B_{x_k, r_k}$  and show that  $t' = \mathcal{F}_{x_k}(t)$  is in  $B_{x_{k+1}, r_{k+1}}$ . Since by (4.31)  $t$  is in the ball  $B_{x_k, \rho(x_k)}$  in standard metric, the estimates in Lemma 4.1 hold for any  $k$ . In particular, by (2), (5)

$$(4.32) \quad \|D_t^{(1)} \mathcal{F}_{x_k}\|_{x_{k+1} \leftarrow x_k} \leq e^{\chi_\ell + 2\varepsilon} \quad \text{and} \quad \|\mathcal{F}_{x_k}(t)\|_{x_{k+1}} \leq e^{\chi_\ell + 2\varepsilon} \|t\|_{x_k} < \|t\|_{x_k},$$

the last since  $\chi_\ell + 2\varepsilon < 0$ , which also yields

$$(4.33) \quad \|t'\|_{x_{k+1}} = \|\mathcal{F}_{x_k}(t)\|_{x_{k+1}} \leq e^{\chi_\ell + 2\varepsilon} r e^{-2Lk\varepsilon} \leq r e^{-2L(k+1)\varepsilon} = r_{k+1},$$

since by the choice of  $\varepsilon$  we have

$$(4.34) \quad \chi_\ell + 2L\varepsilon + 2\varepsilon \leq 0.$$

Estimating  $t'' = R_{k+1}(t')$  using (4.25) and (4.27) we obtain that for any  $R \in B^x(D)$

$$(4.35) \quad \|t''\|_{x_{k+1}} \leq \|t'\|_x \|D^{(N)} R_{k+1}\|_{x_{k+1}, \alpha} \leq r_{k+1} e^{L(k+1)\varepsilon} \|\bar{R}\|_{\mathcal{C}^x} \leq r e^{-L(k+1)\varepsilon} D < \rho(x_{k+1})$$

by (4.31), provided that  $rD < \rho(x)$ .

Now we will show that  $T^x$  is a contraction on  $B^x(D)$ . For this we first estimate  $\|D^{(N)}(T(\bar{R}))_k\|_{x_k, \alpha}$ . We consider

$$(4.36) \quad \begin{aligned} D_t^{(N)}(T^x(R))_k &= D_t^{(N)}((\mathcal{P}_{x_k})^{-1} \circ R_{k+1} \circ \mathcal{F}_{x_k}) = \\ &= D_{t''}^{(1)}(\mathcal{P}_{x_k})^{-1} \circ D_{t'}^{(N)} R_{k+1} \circ D_t^{(1)} \mathcal{F}_{x_k} + J, \end{aligned}$$

where we again denoted  $t' = \mathcal{F}_{x_k}(t)$  and  $t'' = R_{k+1}(t')$ , and where  $J$  consists of a fixed number of terms of the type

$$D_{t''}^{(i)}(\mathcal{P}_{x_k})^{-1} \circ D_{t'}^{(j)} R_{k+1} \circ D_t^{(m)} \mathcal{F}_{x_k}, \quad ijm = N, \quad j < N.$$

Their norm can be estimated using (3.9) as

$$\|D_{t''}^{(i)}(\mathcal{P}_{x_k})^{-1}\|_{x_k \leftarrow x_{k+1}} \cdot \|D_{t'}^{(j)} R_{k+1}\|_{x_{k+1} \leftarrow x_{k+1}}^i \cdot \|D_t^{(m)} \mathcal{F}_{x_k}\|_{x_{k+1} \leftarrow x_k}^{ij}.$$

For the last term we have  $\|D_t^{(m)} \mathcal{F}_{x_k}\|_{x_{k+1} \leftarrow x_k} \leq K(x)C(x)e^{(2k+1)\varepsilon}$  by (4.17). For the middle term we have by (4.26) and (4.32) as  $j \leq N - 1$

$$\|D_{t'}^{(j)} R_{k+1}\|_{x_{k+1} \leftarrow x_{k+1}} \leq \|t'\|_{x_{k+1}}^{1+\alpha} \cdot \|D^{(N)} R_{k+1}\|_{x_{k+1}, \alpha} < \|t\|_{x_k}^{1+\alpha} \cdot \|\bar{R}\|_{\mathcal{C}^x} e^{L(k+1)\varepsilon}.$$

For the first term, since  $t'' \in B_{x_{k+1}, \rho(x_{k+1})}$  by (4.35), we use (4.29) to get

$$(4.37) \quad \|D_{t''}^{(i)}(\mathcal{P}_{x_k})^{-1}\|_{x_k \leftarrow x_{k+1}} \leq c_2(x) e^{Mk\varepsilon},$$

for all  $k$  and  $1 \leq i \leq d$ . Therefore, with  $M'' = M + 2 + L$

$$(4.38) \quad \|J\| < c_4(x) e^{(Mk+L(k+1)+2k+1)\varepsilon} \|t\|_{x_k}^{1+\alpha} \|\bar{R}\|_{\mathcal{C}^x} < c_4(x) e^{M''\varepsilon(k+1)} \|\bar{R}\|_{\mathcal{C}^x} r_k \|t\|_{x_k}^\alpha.$$

Now we estimate the main term in (4.36). As we observed before, estimates in Lemma 4.1 apply to  $t, t', t''$ . In particular, we use (4.32) for  $\mathcal{F}$  term. For the  $\mathcal{P}$  term we claim that

$$(4.39) \quad \|D_{t''}^{(1)}(\mathcal{P}_{x_k})^{-1}\|_{x_k \leftarrow x_{k+1}} \leq e^{\chi_1+2\varepsilon}.$$

This follows from

$$\|D_0^{(1)}(\mathcal{P}_{x_k})^{-1}\|_{x_k \leftarrow x_{k+1}} = \|F_{x_k}^{-1}\|_{x_k \leftarrow x_{k+1}} \leq e^{\chi_1+\varepsilon}$$

similarly to (4.5) in Lemma 4.1. Indeed, if  $d = 1$  then this follows as  $D^{(1)}(\mathcal{P}_{x_k})$  is constant. If  $N \geq 2$  then the Lipschitz constant of  $D^{(1)}(\mathcal{P}_{x_k})^{-1}$  is at most  $c_2(x) e^{Mk\varepsilon}$  by (4.29), so using (4.35) we obtain as  $M < L$

$$\|D_{t''}^{(1)}(\mathcal{P}_{x_k})^{-1} - D_0^{(1)}(\mathcal{P}_{x_k})^{-1}\|_{x_k \leftarrow x_{k+1}} \leq c_2(x) e^{Mk\varepsilon} \|t''\|_{x_{k+1}} \leq c_2(x) r D e^{(M-L)(k+1)\varepsilon} < \varepsilon e^{\chi_1}$$

provided that  $r < e^{\chi_1} (c_2(x) D)^{-1}$ .

We conclude using (4.39), (4.32), and (4.33) that

$$(4.40) \quad \begin{aligned} & \|D_{t''}^{(1)}(\mathcal{P}_{x_k})^{-1} \circ D_{t'}^{(N)} R_{k+1} \circ D_t^{(1)} \mathcal{F}_{x_k}\|_{x_k \leftarrow x_k} \leq \\ & \leq \|D_{t''}^{(1)}(\mathcal{P}_{x_k})^{-1}\|_{x_k \leftarrow x_{k+1}} \cdot \|D^{(N)} R_{k+1}\|_{x_{k+1}, \alpha} \|t'\|_{x_{k+1}}^\alpha \cdot \|D_t^{(1)} \mathcal{F}_{x_k}\|_{x_{k+1} \leftarrow x_k}^N \leq \\ & \leq e^{-\chi_1+2\varepsilon} \cdot \|\bar{R}\|_{\mathcal{C}^x} e^{L(k+1)\varepsilon} e^{\alpha(\chi_\ell+2\varepsilon)} \|t\|_{x_k}^\alpha \cdot e^{N(\chi_\ell+2\varepsilon)} = e^{-\nu+L'\varepsilon} \|t\|_{x_k}^\alpha \|\bar{R}\|_{\mathcal{C}^x} e^{Lk\varepsilon}, \end{aligned}$$

where  $\nu = -(N + \alpha)\chi_\ell + \chi_1 > 0$  and  $L' = 2 + L + 2(N + \alpha)$ . Provided that

$$(4.41) \quad \varepsilon \leq \varepsilon_0 \leq \nu/(2L')$$

we obtain that  $e^{-\nu+L'\varepsilon} \leq e^{-\nu/2} = 1 - 2\theta$  where we defined

$$(4.42) \quad \theta = (1 - e^{-\nu/2})/2 > 0.$$

Combining the estimates (4.38) and (4.40) we get for  $\bar{R} \in B^x(D)$

$$\|D_t^{(N)}(T^x(\bar{R}))_k\|_{x_k \leftarrow x_k} \leq \|t\|_{x_k}^\alpha \|\bar{R}\|_{\mathcal{C}^x} e^{Lk\varepsilon} [1 - 2\theta + c_4(x) r_k e^{\varepsilon(M''(k+1)-Lk)}].$$

Since  $r_k = r e^{-2Lk\varepsilon}$  and  $3L > M''$  we see that for all  $k \geq 0$

$$c_4(x) r_k e^{\varepsilon(M''(k+1)-Lk)} \leq c_4(x) r e^{\varepsilon(M''(k+1)-3Lk)} \leq c_4(x) r e^{\varepsilon M''} \leq \theta$$

if we choose  $r$  satisfying  $r \leq \theta/(c_4(x) e^{\varepsilon M''})$  in addition to  $r < \rho(x)/D = \theta\rho(x)/D'$ . Then for all  $\bar{R} \in B^x(D)$  we obtain

$$\begin{aligned} & \|D^{(N)}(T^x(\bar{R}))_k\|_{x_k, \alpha} \leq (1 - \theta) \|\bar{R}\|_{\mathcal{C}^x} e^{Lk\varepsilon} \quad \text{and} \\ & \|T^x(\bar{R})\|_{\mathcal{C}^x} = \sup_k e^{-Lk\varepsilon} \|D^{(N)}(T^x(\bar{R}))_k\|_{x_k, \alpha} \leq (1 - \theta) \|\bar{R}\|_{\mathcal{C}^x}. \end{aligned}$$



Since  $\|\bar{\Delta}^N\|_{C^x} \leq D' = \theta D$  the operator  $\tilde{T}^x(\bar{R}) = T(\bar{R}) + \bar{\Delta}^N$  is also a contraction and preserves  $B^x(D)$ . Thus  $\tilde{T}^x$  has a unique fixed point  $\bar{R}^x \in B^x(D)$  which depends measurably on  $x$ . As in the construction of Taylor coefficients, the uniqueness implies that  $(R^x)_0$  is  $L\varepsilon$ -tempered and solves the equations (4.23) and (4.22). Thus the measurable family of coordinate changes  $\mathcal{H}_x = \mathcal{H}_x^N + (R^x)_0$ , is also  $L\varepsilon$ -tempered and conjugates  $\mathcal{P}_x$  and  $\mathcal{F}_x$ .

We conclude that the maps  $\mathcal{H}_x$  is a family of  $C^{N,\alpha}$  diffeomorphisms defined on  $B_{x,r(x)}$  which depend measurably on  $x \in \mathcal{M}$  and  $L\varepsilon$ -tempered. Since  $\chi_1 + 2\varepsilon + L\varepsilon < 0$ , we can extend each  $\mathcal{H}_x$  to  $B_{x,\rho(x)}$ . Indeed, by Lemma 4.1 for each  $t \in B_{x,\rho(x)}$  we will have  $\mathcal{F}_x^k(t) \in B_{x_k,r_k}$  for some  $k$ . Then  $\mathcal{H}_x$  is defined uniquely by invariance.

**4.3. Prove of part (2): “uniqueness” of  $\mathcal{H}$ .** This essentially follows from the “uniqueness” of the construction. First we construct inductively coordinate changes  $\mathcal{H}_k = \{\mathcal{H}_{k,x}\}$  for  $k = 1, \dots, N$  with  $\mathcal{H}_1 = \tilde{\mathcal{H}}$ . Consider the Taylor series

$$\mathcal{H}_{1,x}(t) = \sum_{n=1}^{\infty} H_{1,x}^{(n)}(t).$$

By assumption  $H_{1,x}^{(1)} = H_x^{(1)} = \text{Id}$ . Then  $H_{1,x}^{(2)}$  and  $H_x^{(2)}$  satisfy the same equation (4.10) when projected to the factor-bundle  $\mathcal{R}^{(2)}/\mathcal{S}^{(2)}$ . By uniqueness of the solution of (4.10) we obtain that  $H_x^{(2)} = H_{1,x}^{(2)} + \Delta_x^{(2)}$ , where  $\Delta_x^{(2)} \in \mathcal{S}_x^{(2)}$ , and then the polynomial  $\text{Id} + \Delta_x^{(2)}$  is in  $G_\chi$ . Now we consider the coordinate change  $\mathcal{H}_{2,x} = (\text{Id} + \Delta_x^{(2)}) \circ \mathcal{H}_{1,x}$ , which conjugates  $\mathcal{F}$  to a new normal form

$$\mathcal{P}_{2,x} = (\text{Id} + \Delta_{fx}^{(2)}) \circ \mathcal{P}_{1,x} \circ (\text{Id} + \Delta_x^{(2)})^{-1}$$

which is also of sub-resonance type. By construction  $H_{2,x}^{(2)} = H_{1,x}^{(2)} + \Delta_x^{(2)} = H_x^{(2)}$ , so that  $\mathcal{H}$  and  $\mathcal{H}_2$  have the same Taylor terms up to order two.

Inductively, suppose  $\mathcal{H}_{k-1}$  is constricted so that

$$H_{k-1,x}^{(n)} \text{ are } n^2\varepsilon\text{-tempered for } n = 1, \dots, N, \quad H_x^{(n)} = H_{k-1,x}^{(n)} \text{ for } n = 1, \dots, k-1,$$

and the corresponding normal form  $\mathcal{P}_{k-1,x}$  is of sub-resonance type. It follows that  $\mathcal{P}$  and  $\mathcal{P}_{k-1}$  have the same terms up to order  $k-1$ . Hence  $H_{k-1,x}^{(k)}$  and  $H_x^{(k)}$  satisfy the same equation (4.10) when projected to the factor-bundle  $\mathcal{R}^{(k)}/\mathcal{S}^{(k)}$ . Indeed, the  $Q$  term defined by (4.9) is composed only of  $F^{(i)}$  and terms  $H^{(i)}$  and  $P^{(i)}$  with  $1 < i < k$ , which are the same for  $\mathcal{H}_{k-1}$  and  $\mathcal{H}$ . By uniqueness we obtain that

$$H_x^{(k)} = H_{k-1,x}^{(k)} + \Delta_x^{(k)}, \text{ where } \Delta_x^{(k)} \in \mathcal{S}_x^{(k)}.$$

Then the coordinate change  $\mathcal{H}_{k,x} = (\text{Id} + \Delta_x^{(k)}) \circ \mathcal{H}_{k-1,x}$  has the same Taylor terms as  $\mathcal{H}$  up to order  $k$  and, since the polynomial  $\text{Id} + \Delta_x^{(k)}$  is in  $G_\chi$ ,  $\mathcal{H}_k$  conjugates  $\mathcal{F}$  to a sub-resonance normal form  $\mathcal{P}_{k,x} = (\text{Id} + \Delta_{fx}^{(k)}) \circ \mathcal{P}_{k-1,x} \circ (\text{Id} + \Delta_x^{(k)})^{-1}$ . To complete the

inductive step we need to show that  $\|H_{k,x}^{(n)}\|$  is  $n^2\varepsilon$ -tempered. It suffices to show this for  $\|R^{(n)}\|$  where  $R = \Delta_x^{(k)} \circ \mathcal{H}_{k-1,x}$ . Since  $\Delta_x^{(k)}$  is homogeneous of degree  $k$ , we have for  $j = n/k$

$$\|R^{(n)}\| = \|\Delta_x^{(k)} \circ \mathcal{H}_{k-1,x}^{(j)}\| \leq \|\Delta_x^{(k)}\| \cdot \|\mathcal{H}_{k-1,x}^{(j)}\|^k,$$

which is  $(k^2 + j^2k)\varepsilon$ -tempered by the inductive assumption and the definition of  $\Delta_x^{(k)}$ . Since  $j \leq n/2$  as  $k \geq 2$  we get  $j^2k = jn \leq n^2/2$ . If also  $j \geq 2$  we have  $k^2 \leq n^2/4$  and we obtain  $n^2\varepsilon$ -temperedness. If  $j = 1$ , we have  $R^{(k)} = \Delta^{(k)}$ , which is also  $k^2\varepsilon$ -tempered.

Thus in  $N$  steps we obtain the coordinate change

$$\mathcal{H}_{N,x} = G_x \circ \tilde{\mathcal{H}}_x, \quad \text{where } G_x = (\text{Id} + \Delta_x^{(N)}) \circ \cdots \circ (\text{Id} + \Delta_x^{(2)}) \in G_\chi,$$

which has the same Taylor terms at 0 as  $\mathcal{H}$  up to order  $N$ . In fact, for  $n > d$  we have  $\mathcal{S}^{(n)} = 0$  and hence  $\Delta^{(n)} = 0$ , so that  $\mathcal{H}_N = \mathcal{H}_d$ . Now we show that  $\mathcal{H} = \mathcal{H}_N$ , which also proves the last statement in part (2) of the theorem. The equality follows from the uniqueness in the final step of the construction. Indeed the difference  $R = \mathcal{H} - \mathcal{H}_N$  has zero derivatives up to order  $N$  at the zero section and satisfies  $R = T(R)$  for the operator  $T$  from (4.21). Hence  $R = 0$  by uniqueness of the fixed point in the appropriate space  $\mathcal{C}_{r,x}$  on which  $T$  induces a contraction. To ensure that the sequence  $(R_{x_k})$  is in  $\mathcal{C}_{r,x}$  we need estimate temperedness of  $\alpha$ -Hölder constant at 0 for  $\mathcal{H}_N^{(N)}$ . As above one can see that all terms in the polynomial  $G_x$  are  $N^2\varepsilon$ -tempered. Then using Lemma 4.4 and the assumption on  $\tilde{\mathcal{H}}$  we obtain that  $\|\mathcal{H}_{N,x}\|_{C^{N,\alpha}}$  is  $\tilde{L}\varepsilon$ -tempered for  $\tilde{L} = (N^2 + NL) < (N+1)L$ . Hence  $(R_{x_k})$  is in  $\mathcal{C}_{r,x}$  with  $\tilde{L}$  in place of  $L$ , on which  $T$  is a contraction if  $\varepsilon < \varepsilon_1 = \varepsilon_0/(N+1)$ , as the inequalities (4.41) and (4.34) are satisfied. Thus  $(R_{x_k}) = 0$  and extending by invariance, as (4.34) is satisfied, we conclude that  $R_x = 0$  on  $B_{x,\rho(x)}$ , and so  $\mathcal{H} = \mathcal{H}_N$ .

**4.4. Proof of Corollary 2.4.** By part (2) of Theorem 2.3, if we fix a choice of Taylor polynomials of degree  $d$  for  $\mathcal{H}_x$ , then the family  $\mathcal{H}_x$  is unique. Then for each  $N > d$  we can do the construction in part (1) with this fixed choice of Taylor polynomials and obtain the family of  $C^N$  diffeomorphisms  $\mathcal{H}_x$ . By uniqueness, all these families coincide and hence  $\mathcal{H}_x$  are  $C^\infty$  diffeomorphisms.

**4.5. Proof of part (3): Centralizer of  $\mathcal{H}$ .** First we prove that the derivative at zero section  $\Gamma_x = D_0\mathcal{G}_x$  is sub-resonance. This is equivalent to the fact that  $\Gamma_x$  preserves the fast flag associated with the Lyapunov splitting. Suppose to the contrary that for some  $x \in \Lambda$  and some  $i < j$  we have a vector  $t$  in  $E_x^i$  such that  $t' = \Gamma_x(t)$  has nonzero component  $t'_j$  in  $E_{gx}^j$ . Then

$$\|(F_{gx}^n \circ \Gamma_x)(t)\|_{f^n gx} \geq \|F_{gx}^n(t'_j)\|_{f^n gx} \geq e^{(\chi_j - \varepsilon)n} \|t'_j\|_{gx}$$

and on the other hand

$$\|(F_{gx}^n \circ \Gamma_x)(t)\|_{f^n gx} = \|\Gamma_{f^n x}(F_x^n t)\|_{gf^n x} \leq \|\Gamma_{f^n x}\|_{gf^n x \leftarrow f^n x} \cdot e^{(\chi_i + \varepsilon)n} \|t\|_x \leq Ce^{(\chi_i + 3\varepsilon)n}$$

which is impossible for large  $n$  since  $\varepsilon$  is small. Here we used that the  $C^{N,\alpha}$  norm  $\|\mathcal{G}_x\|_{C^{N,\alpha},x}$  with respect to the Lyapunov metric on  $\mathcal{E}_x$  is  $2\varepsilon$ -tempered. This follows as in (4.17) since  $\|\mathcal{G}_x\|_{C^{N,\alpha}}$  in standard norm is  $\varepsilon$ -tempered by assumption.

We conclude that  $\Gamma_x$  is sub-resonance for each  $x \in \Lambda$ . Now we consider a new family of coordinate changes

$$\tilde{\mathcal{H}}_x = \Gamma_x^{-1} \circ \mathcal{H}_{gx} \circ \mathcal{G}_x$$

which also satisfies  $\tilde{\mathcal{H}}_x(0) = 0$  and  $D_0 \tilde{\mathcal{H}}_x = \text{Id}$ . A direct calculation shows that

$$\begin{aligned} \tilde{\mathcal{H}}_{fx} \circ \mathcal{F}_x \circ \tilde{\mathcal{H}}_x &= \Gamma_{fx}^{-1} \circ \mathcal{H}_{fgx} \circ \mathcal{G}_{fx} \circ \mathcal{F}_x \circ \mathcal{G}_x^{-1} \circ \mathcal{H}_{gx}^{-1} \circ \Gamma_x = \\ &= \Gamma_{fx}^{-1} \circ \mathcal{H}_{fgx} \circ \mathcal{F}_{gx} \circ \mathcal{H}_{gx}^{-1} \circ \Gamma_x = \Gamma_{fx}^{-1} \circ \mathcal{P}_{gx} \circ \Gamma_x = \tilde{\mathcal{P}}_x, \end{aligned}$$

where  $\tilde{\mathcal{P}}_x$  is a sub-resonance polynomial as a product of sub-resonance polynomials. Now we would like to apply the uniqueness part of the theorem, which would give  $\tilde{\mathcal{H}}_x = G_x \mathcal{H}_x$  for some tempered function  $G_x \in G_\chi$ . Then it follows from the definition of  $\tilde{\mathcal{H}}_x$  that

$$\mathcal{H}_{gx} \circ \mathcal{G}_x = \Gamma_x \circ \tilde{\mathcal{H}}_x = (\Gamma_x G_x) \circ \mathcal{H}_x$$

so that  $\mathcal{H}_{gx} \circ \mathcal{G}_x \circ \mathcal{H}_x^{-1} = \Gamma_x G_x$ , which is again a sub-resonance polynomial, as claimed.

To complete the proof it remains to show that  $\tilde{\mathcal{H}}_x$  is suitably tempered to obtain uniqueness. As before, we can estimate the Lyapunov norm of the  $n^{\text{th}}$  Taylor term of  $\tilde{\mathcal{H}}_x$  as  $\|\tilde{\mathcal{H}}_x^{(n)}\| = \|\Gamma_x^{-1}\| \circ \|\mathcal{H}_{gx}^{(k)}\| \circ \|\mathcal{G}_x^{(j)}\|^k$  with  $n = kj$  and obtain that it is  $m\varepsilon$ -tempered with  $m \leq 2 + k^2 + 2k < 3n^2$  for  $n \geq 2$ . Since  $\|\mathcal{H}\|_{C^{N,\alpha}}$  is  $L\varepsilon$ -tempered, using Lemma 4.5 below with  $Q = \mathcal{H}$  and  $\mathcal{F} = \mathcal{G}$  we obtain that  $\|\mathcal{H} \circ \mathcal{G}\|_{C^{N,\alpha}}$  is  $(L + 2(N + \alpha))\varepsilon$ -tempered. Then Lemma 4.4 implies that  $\|\tilde{\mathcal{H}}\|_{C^{N,\alpha}}$  is  $3L\varepsilon$ -tempered as  $(2 + L + 2(N + \alpha)) \leq 3L$  (provided that  $L \geq N + 2$ ). So the uniqueness result applies if  $\varepsilon < \varepsilon_* = \varepsilon_1/3 = \varepsilon_0/3(N + 1)$ .

**Lemma 4.5.** *If  $Q$  and  $\mathcal{F}$  are  $C^{N,\alpha}$ , then  $Q \circ \mathcal{F}$  is  $C^{N,\alpha}$  and  $\|Q \circ \mathcal{F}\|_{C^{N,\alpha}} \leq c_N'' \|Q\|_{C^{N,\alpha}} \|\mathcal{F}\|_{C^{N,\alpha}}^{N+\alpha}$ , where  $c_N''$  depends on  $N$  only.*

*Proof.* The proof is the same as in Lemma 4.4 except that, since  $D^{(N)}Q$  is only Hölder, we also need to estimate the  $\alpha$ -Hölder constant at 0 of the term  $D_{\mathcal{F}(t)}^{(N)}Q \circ D_t \mathcal{F}$  in

$$D_t^{(N)}(Q \circ \mathcal{F}) = D_{\mathcal{F}(t)}^{(N)}Q \circ D_t \mathcal{F} + D_{\mathcal{F}(t)}Q \circ D_t^{(N)}\mathcal{F} + \sum_{kj=N, j,k < N} D_{\mathcal{F}(t)}^{(k)}Q \circ D_t^{(j)}\mathcal{F}.$$

$$\begin{aligned} D_{\mathcal{F}(t)}^{(N)}Q \circ D_t \mathcal{F} - D_0^{(N)}Q \circ D_0 \mathcal{F} &= \\ &= (D_{\mathcal{F}(t)}^{(N)}Q - D_0^{(N)}Q) \circ D_t \mathcal{F} + D_0^{(N)}Q \circ D_t^{(N)}\mathcal{F} - D_0^{(N)}Q \circ D_0^{(N)}\mathcal{F} \end{aligned}$$

and its norm can be estimated by

$$\begin{aligned} &\|Q\|_{C^{N,\alpha}} \|\mathcal{F}(t)\|^\alpha \cdot \|D_t \mathcal{F}\|^N + \text{Lip}(D_0^{(N)}Q) \cdot \|D_t \mathcal{F} - D_0 \mathcal{F}\| \leq \\ &\|Q\|_{C^{N,\alpha}} \cdot (\|\mathcal{F}\|_{C^1} \|t\|)^\alpha \cdot \|\mathcal{F}\|_{C^1}^N + c_N' \|D_0^{(N)}Q\| \|\mathcal{F}\|_{C^1}^{N-1} \cdot \|\mathcal{F}\|_{C^{1,\alpha}} \cdot \|t\|^\alpha \leq \\ &\|t\|^\alpha (\|Q\|_{C^{N,\alpha}} \cdot \|\mathcal{F}\|_{C^1}^{N+\alpha} + c_N' \|Q\|_{C^N} \cdot \|\mathcal{F}\|_{C^1}^{N-1} \cdot \|\mathcal{F}\|_{C^{1,\alpha}}). \end{aligned}$$

Here we estimated the Lipschitz constant  $Lip(D_0^{(N)}Q)$  of the homogeneous polynomial  $N$ -form  $D_0^{(N)}Q$  on a ball of radius  $R = \|\mathcal{F}\|_{C^1}$  by the supremum of its derivative on that ball, which is a homogeneous polynomial  $(N-1)$ -form whose norm can be estimated by  $\|D_0^{(N)}Q\|$  with some constant  $c'_N$  depending on  $N$  only.

So the  $\alpha$ -Hölder constant at 0 of  $D_{\mathcal{F}(t)}^{(N)}Q \circ D_t\mathcal{F}$  is estimated by

$$\|Q\|_{C^{N,\alpha}}(\|\mathcal{F}\|_{C^1}^{N+\alpha} + c'_N\|\mathcal{F}\|_{C^{1,\alpha}}^N) \leq (c'_N + 1)\|Q\|_{C^{N,\alpha}}\|\mathcal{F}\|_{C^{N,\alpha}}^{N+\alpha}.$$

We conclude as in Lemma 4.4 that  $\|Q \circ \mathcal{F}\|_{C^{N,\alpha}} \leq c''_N\|Q\|_{C^{N,\alpha}}\|\mathcal{F}\|_{C^{N,\alpha}}^{N+\alpha}$ .

□

This completes the proof of Theorem 2.3.

□

## 5. PROOF OF THEOREM 2.5

**5.1. Proof of (i), (ii), (iii), (v).** We will apply Theorem 2.3. First we note that the integrability condition for the derivative in Theorem 2.3 was used in the proof only to obtain the Lyapunov splitting and the Lyapunov metric. So while the restriction  $Df|_{\mathcal{E}}$  may not satisfy this integrability condition, the Lyapunov splitting and the Lyapunov metric are obtained in this case from the results for the full differential  $Df$ .

The centralizer part (v) will follow directly from (3) of Theorem 2.3 since  $X' = \bigcap_{n \in \mathbb{Z}} g^n(X)$  is the desired invariant set of full measure as  $g$  preserves the measure class of  $\mu$ . Moreover,  $g(W_x) = W_{gx}$  since  $g$  is a diffeomorphism commuting with  $f$ , so that  $X'$  is also saturated by the stable manifolds.

Parts (i), (ii), (iii) essentially follow from Theorem 2.3, which is formulated so as to apply to this setting. First we consider the regular set  $\Lambda$ . We fix a family of local (strong) stable manifolds  $W_{x,r(x)}$  for  $x \in \Lambda$  of sufficiently small size  $r(x)$ . Identifying  $W_{x,r(x)}$  by an exponential map with a neighborhood of 0 in  $\mathcal{E}_x$  we obtain the extension  $\mathcal{F} = \{\mathcal{F}_x\}$  of  $f$ . Then the properties of local stable manifolds ensure that  $\mathcal{F}$  satisfies the assumptions of Theorem 2.3. Indeed, they are given by  $C^{N,\alpha}$  embeddings so that the  $C^{N,\alpha}$  norm and  $1/r(x)$  are  $\varepsilon$ -tempered for any  $\varepsilon > 0$  (see [BP] for a general reference and [KtR15, Theorem 5] for a convenient statement of the stable manifold theorem). Hence Theorem 2.3 yields existence of the desired family of local diffeomorphisms  $\mathcal{H}_x$ ,  $x \in \Lambda$ , which can be uniquely extended to global diffeomorphisms by invariance.

Now we define  $X = \bigcup_{x \in \Lambda} W_x$  and explain the construction of  $\mathcal{H}_y$  for any  $y \in X$ . By iterating it forward we may assume that  $y \in W_{x,r(x)}$ . While the individual Lyapunov spaces  $\mathcal{E}^i$  may not be defined for all points  $y \in W_{x,r(x)}$ , the flag  $\mathcal{V}$  of fast subspaces

$$(5.1) \quad \mathcal{E}_x^1 = \mathcal{V}_x^1 \subset \mathcal{V}_x^2 \subset \dots \subset \mathcal{V}_x^l = \mathcal{E}_x, \quad \text{where } \mathcal{V}_x^i = \mathcal{E}_x^1 \oplus \dots \oplus \mathcal{E}_x^i,$$

is defined for each  $\mathcal{E}_y = T_y W_{x,r(x)}$ . Moreover, the subspaces  $\mathcal{V}_y^i$  depend Hölder continuously, and in fact  $C^{N-1,\alpha}$ , on  $y$  along  $W_x$  [R79, Theorem 6.3].

The key observation is that the notion of sub-resonance polynomial depends only on the fast flag  $\mathcal{V}$  [KS15, Proposition 3.2], not on the individual Lyapunov spaces  $\mathcal{E}^i$ , and

thus is well-defined for  $\mathcal{E}_y$ . Then the sub-bundle  $\mathcal{S}^{(n)}$  of sub-resonance polynomials of degree  $n$  is well-defined, invariant under  $Df$ , and Hölder continuous in  $y$  along  $W$ , and hence so is the factor bundle  $\mathcal{R}^{(n)}/\mathcal{S}^{(n)}$ . Then for each  $y \in W_{x,r(x)}$  we can define  $\mathcal{H}_y$  using the construction in Theorem 2.3. Indeed, first we constructed the Taylor term of degree  $n$  using the contraction  $\tilde{\Phi}$  on the bundle  $\mathcal{R}^{(n)}/\mathcal{S}^{(n)}$  from Lemma 4.3 with linear part estimated as  $\|\Phi_x(R)\|_{\varepsilon,x} \leq e^{\lambda+(n+1)\varepsilon} \|R\|_{\varepsilon,fx}$ . Then  $\Phi_y$ , the corresponding map at  $y$  is Hölder close to  $\Phi_x$ . Using the Lyapunov norm at  $x$  as the reference norm, we obtain that  $\Phi_y$  is also a contraction with similar estimate for all  $y \in W_{x,r(x)}$  provided that  $r(x)$  is sufficiently small. Since  $f^k y \in W_{f^k x, r(f^k x)}$  by the contraction property of  $W_{x,r(x)}$ , the closeness persists along the forward trajectory. This argument is similar to the proof of Lemma 4.1. Then we obtain that the operator  $\tilde{\Phi}_y$  on the sequence space is also a contraction. Thus we can define  $\bar{\mathcal{H}}_y^{(n)}$  as before using the unique fixed point in the space of sequences. The last step of the construction can be carried out similarly as it involves only the estimates of the derivatives on the full space  $\mathcal{E}$  and does not depend on the splitting.

**Remark 5.1.** Any measurable choice of transversals  $\tilde{\mathcal{E}}^i$  to  $\mathcal{V}^{i-1}$  inside  $\mathcal{V}^i$ ,  $i = 2, \dots, \ell$ , yields a transversal  $\tilde{\mathcal{N}}^{(n)}$  to  $\mathcal{S}^{(n)}$  inside  $\mathcal{R}^{(n)}$ . The latter gives a preferred choice of the lift. The fixed point of the contraction  $\bar{\mathcal{H}}_y^{(n)}$  depends Hölder continuously (and even smoothly by appropriate  $C^r$  section theorem as in [KS15]) on  $y$  along  $W_{x,r(x)}$  if the same holds for the data  $\tilde{Q}$  obtained in the previous step of the construction. To complete the inductive step we need a Hölder lift  $\bar{\mathcal{H}}_y^{(n)}$  to  $\mathcal{R}^{(n)}$ . If there is a consistent choice which is Hölder on the full leaves of  $W$ , then we can obtain a family  $\{\mathcal{H}_x\}$  which is Hölder along the leaves of  $W$ . In contrast to the uniform setting of [KS15], it is not clear that such a choice exists. However, this can be done locally on  $W_{x,r(x)}$ , so then one can fix a Ledrappier-Young partition subordinate to the leaves of  $W$  and obtain Hölder continuity of  $\mathcal{H}_x$  on each element.

**5.2. Consistency of the fast foliations.** The leaf  $W_x$  is subfoliated by unique foliations  $U^k$  tangent to  $\mathcal{V}_y^k$ . We denote by  $\bar{W}^k$  the corresponding foliations of  $\mathcal{E}_x$  obtained by the identification  $\mathcal{H}_x : W_x \rightarrow \mathcal{E}_x$ . Thus we obtain the foliations  $\bar{W}^k$  of  $\mathcal{E}$  which are invariant under the polynomial extension  $\mathcal{P}$ . Since the maps  $\mathcal{H}_x$  are diffeomorphisms,  $\bar{W}^k$  are also the unique fast foliations with the same contraction rates. They are characterized, for any  $\varepsilon$  sufficiently small so that  $\chi_k + \varepsilon < \chi_{k+1}$ , by

$$\text{for } y, z \in \mathcal{E}_x \quad z \in \bar{W}^k(y) \Leftrightarrow \text{dist}(\mathcal{P}_x^n(y), \mathcal{P}_x^n(z)) \leq C e^{n(\chi_k + \varepsilon)} \text{ for all } n \in \mathbb{N}.$$

It follows from Definition 2.2 that sub-resonance polynomials  $R \in \mathcal{S}_{x,y}$  are *block triangular* in the sense that  $\mathcal{E}^i$  component does not depend on  $\mathcal{E}^j$  components for  $j < i$  or, equivalently, it maps map the subspaces  $\mathcal{V}_x^i$  of fast flag in  $\mathcal{E}_x$  to those in  $\mathcal{E}_y$ .

It is easy to see that all derivatives of a sub-resonance polynomial are sub-resonance polynomials. In particular, the derivative  $D_y \mathcal{P}_x$  at *any* point  $y \in \mathcal{E}_x$  is sub-resonance

and hence is block triangular. Thus it maps subspaces parallel to  $\mathcal{V}_x^k$  to subspaces parallel to  $\mathcal{V}_{f_x}^k$ . Hence the foliation of  $\mathcal{E}$  by subspaces parallel to  $\mathcal{V}_x^k$  in  $\mathcal{E}_x$  is invariant under the extension  $\mathcal{P}$  and hence coincides with  $\bar{W}^k$  by uniqueness of the fast foliation.

**Remark 5.2.** *This implies that the fast subfoliations  $U^k$  are as smooth along the leaf  $W_x$  as the diffeomorphism  $\mathcal{H}_x$  which maps them to linear subfoliations of  $\mathcal{E}_x$ .*

It follows that for any  $x \in \mathcal{M}$  and any  $y \in W_x$  the diffeomorphism

$$(5.2) \quad \mathcal{G}_{x,y} := \mathcal{H}_y \circ \mathcal{H}_x^{-1} : \mathcal{E}_x \rightarrow \mathcal{E}_y$$

maps the fast flag of linear foliations of  $\mathcal{E}_x$  to that of  $\mathcal{E}_y$ . In particular, the same holds for its derivative  $D_0 \mathcal{G}_{x,y} = D_x \mathcal{H}_y : \mathcal{E}_x \rightarrow \mathcal{E}_y$  and we conclude that  $D_0 \mathcal{G}_{x,y}$  is block triangular and thus is a sub-resonance linear map.

**5.3. Proof of (iv): Consistency of normal form coordinates.** We need to show that the map  $\mathcal{G}_{x,y}$  in (5.2) is a sub-resonance polynomial map for all  $x \in X$  and  $y \in W_x$ . It suffices to consider  $x \in \Lambda$  and, using invariance, we may assume that  $y \in W_x$  is sufficiently close to  $x$ . First we note that

$$\mathcal{G}_{x,y}(0) = \mathcal{H}_y(x) =: \bar{x} \in \mathcal{E}_y \quad \text{and} \quad D_0 \mathcal{G}_{x,y} = D_x \mathcal{H}_y.$$

Since  $\mathcal{H}_{f^n x}^{-1} \circ \mathcal{P}_x^n \circ \mathcal{H}_x = f^n = \mathcal{H}_{f^n y}^{-1} \circ \mathcal{P}_y^n \circ \mathcal{H}_y$  we obtain that

$$\mathcal{H}_{f^n y} \circ \mathcal{H}_{f^n x}^{-1} \circ \mathcal{P}_x^n = \mathcal{H}_{f^n y} \circ f^n \circ \mathcal{H}_x^{-1} = \mathcal{P}_y^n \circ \mathcal{H}_y \circ \mathcal{H}_x^{-1} \quad \text{and hence}$$

$$(5.3) \quad \mathcal{G}_{f^n x, f^n y} \circ \mathcal{P}_x^n = \mathcal{P}_y^n \circ \mathcal{G}_{x,y}.$$

Now we consider the Taylor polynomial for  $\mathcal{G}_{x,y} : \mathcal{E}_x \rightarrow \mathcal{E}_y$  at  $t = 0 \in \mathcal{E}_x$ :

$$\mathcal{G}_{x,y}(t) \sim G_{x,y}(t) = \bar{x} + \sum_{m=1}^N G_{x,y}^{(m)}(t).$$

Our first goal is to show that all its terms are sub-resonance polynomials. We proved in Section 5.2 that the first derivative  $G_{x,y}^{(1)} = D_x \mathcal{H}_y$  is a sub-resonance linear map.

Inductively, we assume that  $G_{x,y}^{(m)}$  has only sub-resonance terms for  $m = 1, \dots, k-1$  and show that the same holds for  $G_{x,y}^{(k)}$ . Suppose for the contrary that  $G_{x,y}^{(k)}$  is not a sub-resonance polynomial and consider order  $k$  terms in the Taylor polynomial at  $0 \in \mathcal{E}_x$  for (5.3). Taylor polynomial for  $\mathcal{P}_x^n$  at  $0 \in \mathcal{E}_x$  coincides with  $\mathcal{P}_x^n(t) = \sum_{m=1}^d P_x^{(m)}(t)$ . We also consider the Taylor polynomial for  $\mathcal{P}_y^n$  at  $\mathcal{G}_{x,y}(0) = \bar{x} \in \mathcal{E}_y$

$$\mathcal{P}_y^n(z) = \bar{x}_n + \sum_{m=1}^d Q_y^{(m)}(z - \bar{x}), \quad \text{where } \bar{x}_n = \mathcal{P}_y^n(\bar{x}).$$

All terms  $Q^{(m)}$  are sub-resonance as the derivatives of a sub-resonance polynomial. Consider the Taylor polynomial for

$$\mathcal{G}_{f^n x, f^n y}(t) \sim G_{f^n x, f^n y}(t) = \bar{x}_n + \sum_{m=1}^N G_{f^n x, f^n y}^{(m)}(t).$$

Now we obtain from (5.3) the coincidence of the terms up to degree  $N$  in

$$\bar{x}_n + \sum_{j=1}^N G_{f^n x, f^n y}^{(j)} \left( \sum_{m=1}^d P_x^{(m)}(t) \right) \sim \bar{x}_n + \sum_{m=1}^d Q_y^{(m)} \left( \sum_{j=1}^N G_{x, y}^{(j)}(t) \right).$$

Since any composition of sub-resonance polynomials is again sub-resonance, the inductive assumption gives that all terms of order  $k$  in the above equation must be sub-resonance polynomials except for

$$G_{f^n x, f^n y}^{(k)}(P_x^{(1)}(t)) \quad \text{and} \quad Q_y^{(1)}(G_{x, y}^{(k)}(t)).$$

Multiplying these terms on the left by sub-resonance linear map  $(D_0 G_{f^n x, f^n y}^{(k)})^{-1} = (D_{f^n x} \mathcal{H}_{f^n y})^{-1}$  and using the fact that  $P_x^{(1)} = F_x^n = Df^n|_{\mathcal{E}_x}$  and

$$Q_y^{(1)} = D_{\bar{x}} \mathcal{P}_y^n = D_{f^n x} \mathcal{H}_{f^n y} \circ F_x^n \circ (D_x \mathcal{H}_y)^{-1}$$

we obtain that the following maps from  $\mathcal{E}_x$  to  $\mathcal{E}_{f^n x}$  agree modulo sub-resonance terms

$$\left( (D_{f^n x} \mathcal{H}_{f^n y})^{-1} \circ G_{f^n x, f^n y}^{(k)} \right) \circ F_x^n \cong F_x^n \circ \left( (D_x \mathcal{H}_y)^{-1} \circ G_{x, y}^{(k)} \right) \quad \text{mod } \mathcal{S}_{x, f^n x}.$$

Since  $x, f^n x \in \Lambda$  and thus the spaces  $\mathcal{E}_x$  and  $\mathcal{E}_{f^n x}$  have Lyapunov splittings we can decompose these polynomial maps into sun-resonance and non sub-resonance terms. Taking non sub-resonance terms on both sides we obtain the equality

$$(5.4) \quad N_{f^n x} \circ F_x^n = F_x^n \circ N_x$$

where  $N_{f^n x}$  and  $N_x$  denote the non sub-resonance terms in  $(D_{f^n x} \mathcal{H}_{f^n y})^{-1} \circ G_{f^n x, f^n y}^{(k)}$  and  $(D_x \mathcal{H}_y)^{-1} \circ G_{x, y}^{(k)}$  respectively. If the latter had only sub-resonance terms then so would  $G_{x, y}^{(k)}$ , contradicting the assumption. Hence  $N_x \neq 0$ . We decompose  $N_x$  into components  $N_x = (N_x^1, \dots, N_x^\ell)$  and let  $i$  be the largest index so that  $N_x^i \neq 0$ , i.e. there exists  $t' \in \mathcal{E}_x$  so that  $z' = N(t')$  has non-zero component in  $\mathcal{E}_y^i$ , which we denote by  $z'_i$ . Then by (3.3) we obtain

$$(5.5) \quad \|F_x^n \circ N_x(t')\|_{f^n x} = \|F_x^n(z')\|_{f^n x} \geq e^{n(\chi_i - \varepsilon)} \|z'_i\|_x.$$

Now we estimate the norm of the  $i$  component of the left-hand side of (5.4) at  $t'$ . For each componet  $t'_j$  of  $t'$  we have  $\|F_x^n(t'_j)\|_{f^n x} \leq e^{n(\chi_j + \varepsilon)} \|t'_j\|_x$  by (3.3). Let  $N_{f^n x}^s$  be a term of homogeneity type  $s = (s_1, \dots, s_\ell)$  in the component  $N_{f^n x}^i$ . Then we obtain as in Lemma 4.2

$$\|N_{f^n x}^s(F_x^n(t'))\|_{f^n x} \leq \|N_{f^n x}\|_{f^n x} \cdot \|t'\|_x^k \cdot e^{n \sum s_j (\chi_j + \varepsilon)}.$$

Since no term in  $N_{f^{n_x}}^i$  is a sub-resonance one, we have  $\chi_i > \sum s_j \chi_j$ . This contradicts (5.4) and (5.5) for large  $n$  if  $\varepsilon$  is sufficiently small since  $\|N_{f^{n_x}}\|_{f^{n_x}}$  is tempered. The latter follows from temperedness of  $G_{f^{n_x}, f^{n_y}}^{(k)}$  and the fact that  $D_{f^{n_x}} \mathcal{H}_{f^{n_y}}$  is Hölder close to the identity and so the norm of its inverse is bounded in Lyapunov metric.

We conclude that for all  $x \in X$  and  $y \in W_x$  the Taylor polynomial  $G_{x,y}$  of  $\mathcal{G}_{x,y}$  contains only sub-resonance terms. Now we will show that  $\mathcal{G}_{y,x}$  coincides with its Taylor polynomial. Again it suffices to consider  $x \in \Lambda$  and  $y \in W_x$  which is sufficiently close to  $x$ . In addition to (5.3) we have the same relation for their Taylor polynomials

$$(5.6) \quad G_{f^{n_y}, f^{n_x}} \circ \mathcal{P}_y^n = \mathcal{P}_x^n \circ G_{y,x}.$$

Indeed, the two sides must have the same terms up to order  $N$ , but these are sub-resonance polynomials and thus have no terms of degree higher than  $d \leq N$ .

Denoting  $\Delta_n = \mathcal{G}_{f^{n_y}, f^{n_x}} - G_{f^{n_y}, f^{n_x}}$  we obtain from (5.3) and (5.6) that

$$(5.7) \quad \Delta_n \circ \mathcal{P}_y^n = \mathcal{P}_x^n \circ \mathcal{G}_{y,x} - \mathcal{P}_x^n \circ G_{y,x}.$$

We denote  $\Delta = \mathcal{G}_{y,x} - G_{y,x} : \mathcal{E}_y \rightarrow \mathcal{E}_x$  and suppose that  $\Delta \neq 0$ . Let  $i$  be the largest index for which the  $i$  component of  $\Delta$  is nonzero. Then there exist arbitrarily small  $t' \in \mathcal{E}_y$  such that the  $i$  component  $z'_i$  of  $z' = \Delta(t')$  is nonzero. Since  $\mathcal{P}_x^n$  is a sub-resonance polynomial, the nonlinear terms in its  $i$  component can depend only on  $j$  components of the input with  $j > i$ , which are the same for  $\mathcal{G}_{y,x}$  and  $G_{y,x}$ . Thus the  $i$  component of the right side of (5.7) is  $F_x^n(z'_i)$  since the linear part of  $\mathcal{P}_x^n$  is  $F_x^n$  and it preserves the Lyapunov splitting. So by (3.3) we can estimate the right side of (5.7)

$$(5.8) \quad \|(\mathcal{P}_x^n \circ \mathcal{G}_{y,x} - \mathcal{P}_x^n \circ G_{y,x})(t')\|_{f^{n_x}} \geq \|F_x^n(z'_i)\|_{f^{n_x}} \geq e^{n(\chi_i - \varepsilon)} \|z'_i\|_x \geq e^{n(\chi_1 - \varepsilon)} \|z'_i\|_x.$$

Now we estimate the left side of (5.7). Since  $\mathcal{G}_{f^{n_y}, f^{n_x}}$  is  $C^{N,\alpha}$  there exists  $C_n$  determined by  $\|\mathcal{G}_{f^{n_x}, f^{n_y}}\|_{C^{N,\alpha}}$  such that

$$(5.9) \quad \|\Delta_n(t)\| \leq C_n \|t\|^{N+\alpha} \quad \text{for all } t \in \mathcal{E}_{f^{n_x}} \text{ with } \|t\| \leq r_n.$$

To estimate  $\mathcal{P}_y^n$  we note that  $D_0 \mathcal{P}_y^n = F_y^n = Df^n|_{\mathcal{E}_y}$  and its norm for  $y$  close to  $x$  can be estimated using Lemma 4.1(3). Then  $\mathcal{P}_y^n$  itself can be estimated as in that lemma:

$$\|\mathcal{P}_y^n(t)\| \leq K e^{n(\chi_\ell + 3\varepsilon)} \|t\|$$

for all sufficiently small  $t \in \mathcal{E}_y$ . Combining this with (5.9) we obtain

$$\|(\Delta_n \circ \mathcal{P}_y^n)(t')\| \leq C_n \|\mathcal{P}_y^n(t')\|^{N+\alpha} \leq C_n (K \|t'\|)^{N+\alpha} e^{n(N+\alpha)(\chi_\ell + 3\varepsilon)}.$$

Now we see that this contradicts (5.7) and (5.8) for large  $n$  if  $\varepsilon$  is sufficiently small. Indeed  $(N + \alpha)\chi_\ell < \chi_1$  while  $C_n$  is tempered and the Lyapunov norm satisfies  $\|u\| \geq K(x) e^{-n\varepsilon} \|u\|_{f^{n_x}}$ . Thus,  $\Delta = 0$ , i.e. the map  $\mathcal{G}_{y,x}$  coincides with its Taylor polynomial.

This completes the proof of Theorem 2.5.  $\square$



**5.4. Proof of Corollary 2.6.** If  $d = 1$  then all sub-resonance polynomials are linear, the maps  $\mathcal{H}_y \circ \mathcal{H}_x^{-1} : \mathcal{E}_x \rightarrow \mathcal{E}_y$  are affine, and the family  $\{\mathcal{H}_x\}_{x \in X}$  is unique by part (2) of Theorem 2.3. If we identify  $W_x$  with  $\mathcal{E}_x$  by  $\mathcal{H}_x$ , then  $\mathcal{H}_y$  for  $y \in W_x$  becomes an affine map  $\mathcal{E}_x \rightarrow T_y \mathcal{E}_x$  with identity differential and  $\mathcal{H}_y(y) = 0$ . Thus it depends  $C^N$  on  $y$  as the coordinate system  $\mathcal{H}_x$  is  $C^N$ .

## REFERENCES

- [BP] L. Barreira and Ya. Pesin. Nonuniformly Hyperbolicity: Dynamics of Systems with Nonzero Lyapunov Exponents. Encyclopedia of Mathematics and Its Applications, **115** Cambridge University Press.
- [BKo] I. U. Bronstein and A. Ya. Kopanskii. Smooth invariant manifolds and normal forms. World Scientific, 1994.
- [F07] Y. Fang. On the rigidity of quasiconformal Anosov flows. Ergodic Theory Dynam. Systems 27 (2007), no. 6, 1773-1802.
- [FFH10] Y. Fang, P. Foulon, and B. Hasselblatt. Zygmund strong foliations in higher dimension. J. Mod. Dyn. 4 (2010), no. 3, 549-569.
- [Fe04] R. Feres. A differential-geometric view of normal forms of contractions. In Modern Dynamical Systems and Applications, Eds.: M. Brin, B. Hasselblatt, Y. Pesin, Cambridge University Press, (2004) 103-121.
- [FKSp11] D. Fisher, B. Kalinin, and R. Spatzier. Totally nonsymplectic Anosov actions on tori and nilmanifolds. Geom. Topol. 15 (2011), no. 1, 191- 216.
- [GoKS11] A. Gogolev, B. Kalinin, and V. Sadovskaya. Local rigidity for Anosov automorphisms. (with Appendix by R. de la Llave) Mathematical Research Letters, 18 (2011), no. 05, 843-858.
- [G02] M. Guysinsky. The theory of non-stationary normal forms. Ergod. Theory Dyn. Syst., **22** (3), (2002), 845-862.
- [GKt98] M. Guysinsky and A. Katok. Normal forms and invariant geometric structures for dynamical systems with invariant contracting foliations. Math. Research Letters **5** (1998), 149-163.
- [KKt01] B. Kalinin and A. Katok. Invariant measures for actions of higher rank abelian groups. Proceedings of Symposia in Pure Mathematics. Volume **69**, (2001), 593-637.
- [KKt07] B. Kalinin and A. Katok. Measure rigidity beyond uniform hyperbolicity: invariant measures for Cartan actions on tori. Journal of Modern Dynamics, Vol. 1 (2007), no. 1, 123-146.
- [KKtR11] B. Kalinin, A. Katok, and F. Rodriguez-Hertz. Nonuniform measure rigidity. Annals of Mathematics 174 (2011), no. 1, 361-400.
- [KtR15] A. Katok and F. Rodriguez-Hertz. Arithmeticity and topology of higher rank actions of Abelian groups. Journal of Modern Dynamics, Vol. 10 (2016) 115-152.
- [KtL91] A. Katok and J. Lewis. Local rigidity for certain groups of toral automorphisms. Israel J. Math. **75** (1991), 203-241.
- [KtSp97] A. Katok and R. Spatzier. Differential rigidity of Anosov actions of higher rank abelian groups and algebraic lattice actions. Tr. Mat. Inst. Steklova 216 (1997), Din. Sist. i Smezhnye Vopr., 292319; translation in Proc. Steklov Inst. Math. 1997, no. 1 (216), 287-314.
- [KS03] B. Kalinin and V. Sadovskaya. On local and global rigidity of quasiconformal Anosov diffeomorphisms. Journal of the Institute of Mathematics of Jussieu (2003) 2 (4), 567-582.
- [KS06] B. Kalinin and V. Sadovskaya. Global rigidity for totally nonsymplectic Anosov  $\mathbb{Z}^k$  actions. Geometry and Topology, vol. 10 (2006), 929-954.
- [KS15] B. Kalinin and V. Sadovskaya. Normal forms on contracting foliations: smoothness and homogeneous structure. To appear in Geometriae Dedicata.

- [LL05] W. Li and K. Lu. Sternberg theorems for random dynamical systems. Communications on Pure and Applied Mathematics, Vol. LVIII (2005), 0941-0988.
- [M16] K. Melnick. Nonstationary smooth geometric structures for contracting measurable cocycles. <http://www.math.umd.edu/~kmelnick/>
- [R79] D. Ruelle. Ergodic theory of differentiable dynamical systems. Publications Mathématiques de l'I.H.É.S. 50 (1979), 27-58.
- [S05] V. Sadovskaya. On uniformly quasiconformal Anosov systems. Math. Research Letters, vol. 12 (2005), no. 3, 425-441.
- [St57] S. Sternberg. Local contractions and a theorem of Poincaré. Amer. J. of Math. 79 (1957), 809-824.

DEPARTMENT OF MATHEMATICS, THE PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK,  
PA 16802, USA.

*E-mail address:* `kalinin@psu.edu`, `sadovskaya@psu.edu`